Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *University of Michigan*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Performance in Eyelid Conditioning Related to Changes in Muscular Tension and Physiological Measures of Emotionality: W. N. RUNQUIST AND K. W. SPENCE 417

Intermittent Primary Reinforcement as a Parameter of Secondary Reinforcement: R. M. KLEIN 423

The Efficiency of Utilization of Visual Information and the Effects of Stress: A. JONES 428

The Perception of Repetition Rate: J. C. STEVENS AND G. M. SHICKMAN 433

Temporal Numerosity: IV. A Comparison of the Major Senses: C. T. WHITE AND P. G. CHEATHAM 441

The von Restorff Isolation Effect: Test of the Intralist Association Assumption: E. SALTZ AND S. E. NEWMAN 445

Intentional and Incidental Learning under High and Low Emotional Drive Levels: D. H. KAUSLER, E. P. TRAPP, AND C. L. BREWER 452

The Role of Past Events in Problem Solving: J. J. GOODNOW, I. RUBINSTEIN, AND B. L. SHANKS 456

Effects of Round Number Preferences in a Guessing Task: BRUCE M. ROSS AND TRYGG ENGEN 462

Sensory Preconditioning: Central Linkage or Response Mediation?: .....S. H. LOVIBOND 469

The Effect of Fixation Conditions on Depth Discrimination Thresholds at Scotopic and Photopic Illuminance Levels: A. LIT 476

Escape Learning as a Function of Amount of Shock Reduction: G. H. BOWER, H. FOWLER, AND M. A. TRAPOLD 482

Spatial Contiguity of Cue, Reward, and Response in Discrimination Learning by Children: J. V. MURPHY AND R. E. MILLER 485

Supplementary Report: Stimulus and Response Meaningfulness (m') in Paired-Associate Learning by Hospitalized Mental Patients: V. J. CIEUTAT 490

American Psychological Association

Vol. 58, No. 6 December 1959

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1959 the publication cost is $30 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
BEATRICE B. HAMLIN, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1959 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 58, NO. 6 DECEMBER, 1959

PERFORMANCE IN EYELID CONDITIONING RELATED TO CHANGES IN MUSCULAR TENSION AND PHYSIOLOGICAL MEASURES OF EMOTIONALITY [1]

W. N. RUNQUIST AND K. W. SPENCE

State University of Iowa

In interpreting some recent studies of classical aversive conditioning use has been made of the basic assumption that the associative factor (H) and the drive factor (D) combine in a multiplicative manner to determine excitatory strength (E), which in turn determines performance. Drive level has been conceptualized in terms of an emotional response (r_e) to the noxious stimulation (UCS) of preceding trials, so that the greater the amplitude of r_e, the higher the drive level (Spence: 1956, 1958). Studies in which performance level in eyelid conditioning has been shown to be a function of UCS intensity (Spence, 1953; Spence & Taylor, 1951) and a function of extreme scores on a questionnaire designed to select *S*s differing in emotionality (Spence & Farber, 1953; Spence & Taylor, 1951; Taylor, 1951) have been offered in support of this theory. In two recent studies (Runquist & Ross: 1958, 1959), emotional *S*s were selected on the basis of either large pulse rate increases *or* skin conductance increases in response to weak noxious stimulation (an air puff to the eye). The performance of these *S*s in eyelid conditioning was compared with that of a group which did not show large increases in these measures. The emotional group performed at a higher level than the nonemotional group. This latter result has been interpreted as supporting the hypothesis that drive level (D) is a function of the intensity of the emotional response (r_e) to noxious stimuli, and as also supporting a hypothesis advanced by Lacey (1956) that the "symptoms" of emotionality are specific to each individual. That is, not all physiological systems are activated by stress or noxious stimulation, but there are individual differences not only in intensity of amplitude of the emotional response but also in the particular autonomic channel through which it discharges.

[1] This research was carried out as a part of a project concerned with the influence of motivation on performance in learning under Contract N9 onr-93802, Project NR-154-107 between the State University of Iowa and the Office of Naval Research. Thanks is given to John Hunter who served as research assistant.

The concept of level of emotional response as underlying generalized drive level (D) in behavior theory (Spence, 1958) corresponds quite closely to the concepts of level of activation and arousal level recently discussed by Duffy (1951), Schlosberg (1954), Lindsley (1951), Hebb (1955), Freeman (1948), and Malmo (1957). These authors refer to arousal level as the intensity dimension of behavior in contrast to its directive aspects. That is, arousal level affects any response being made regardless of what the particular response is. The definition of level of arousal has been in terms of certain physiological measures, such as EEG and autonomic activity (pulse rate, GSR, respiration, etc.). Several recent studies by Malmo and his associates (Bartoshuk, 1955; Malmo, 1957; Malmo & Davis, 1956; Stennett, 1957) have suggested that changes in the level of muscular tension, as measured by muscle action potential (MAP), could also be used to define a level of arousal continuum. In fact, Malmo notes the similarity between arousal level and anxiety as defined by the Manifest Anxiety Scale and suggests the use of physiologically defined arousal level as a possible alternative.

If the emotional responsiveness of an individual is related to MAP changes, then those *Ss* with large MAP changes should show a higher level of conditioning performance than those *Ss* with small MAP changes. Moreover, the conditioning curves should diverge with trials. These predictions follow from the hypothesized relation between emotional responsiveness and drive level, D, and the multiplicative relation between D and H. The primary purpose of this study was to investigate the relation between MAP change and level of performance in eyelid conditioning. A second purpose was to check on results obtained in the previous studies (Runquist & Ross: 1958, 1959) relating conditioning performance to level of autonomic activity when pulse rate and skin conductance are measured during the course of conditioning rather than in a pretest series of UCS presentations as was done in the previous investigations.

Method

Subjects.—The *Ss* were 86 volunteers from an introductory course in psychology. An additional 35 *Ss* were run but were discarded. Apparatus failure accounted for 22, 5 were classified voluntary responders, and 8 gave initial CRs to the light CS.

Apparatus.—The equipment for recording eyeblinks and delivering the UCS, a 1-lb. air puff to the right eye, was the same as that used in previous studies from this laboratory. The CS was the increase in brightness of a 6-cm. circular disc. The duration of the CS was 550 msec. with the UCS occurring 550 msec. after its onset. The air puff was limited to 50 msec. by a solenoid valve controlled by an electronic timer.

The equipment for recording pulse rate and skin resistance was the same as used in a previous study (Runquist & Ross, 1959). EKG electrodes were attached to each forearm, and GSR electrodes to the palm and back of the right hand. The muscle action potentials were recorded with one electrode placed above the sternomastoid muscle about half way up the neck. The ground electrode was placed on the back of the neck. The action potentials were integrated by means of a full wave rectifier, then sent through a low pass filter to a second amplifier. The MAP appears as a low frequency of the order of 10–15 cps, and is recorded by a Brush ink-writing oscillograph. The amplitude of the record was a summation of the energy generated within each period of time controlled by the low pass filter. Amplifiers used push-pull circuits with standard high quality broadcast transformers. An over-all amplification of about one million was obtained with this method.

Procedure.—After all electrodes had been attached, the standard eyelid conditioning procedure employed in this laboratory was used. All *Ss* received three trials with the CS alone, plus one trial with the UCS alone, followed by 80 conditioning trials. The inter-

trial interval was either 15, 20, or 25 sec. in a prearranged order. A verbal ready signal preceded each trial by from 2 to 4 sec., also in a prearranged order. A CR was recorded when a deflection of 1 mm. occurred in the interval 200–500 msec. following the onset of the CS. Measurement of physiological responses was made for a 5-sec. period before the first CS alone trial, on the UCS alone trial, and on trials 10, 20, 30, 40, 50, 60, 70, and 80 of conditioning. On these trials a continuous record of pulse rate, skin conductance, and MAP was obtained for 15 sec. starting 5 sec. before the ready signal. The intertrial interval preceding the test trials was always 25 sec., and the time between ready signal and CS onset on these trials was always 2 sec.

Measurement of physiological responses.—Pulse rate was measured as the speed of two beats at points immediately preceding the onset of the CS and 4 sec. after the UCS. Skin conductance measures were also taken at the same points. The MAP measure was the sum of the magnitude of the deflections during a 500 msec. interval preceding the CS and immediately following the UCS. Measures of all three responses were made at a randomly selected 500 msec. interval while *S* was resting before the experimental procedure was begun.

Results

Autonomic responses.—By selecting samples of emotional and nonemotional *S*s on the basis of the two autonomic measures (pulse rate and skin conductance), a comparison may be made between the present study and two previous ones (Runquist & Ross: 1958, 1959). Two separate analyses were made. In the first analysis pulse rate increase and skin conductance increase in response to the air puff delivered alone on a single trial were used. The emotional *S*s were those who were in the 15% of the sample with the largest pulse rate increase *or* the 15% of the sample with the largest skin conductance increase ($N = 20$). The nonemotional *S*s were the 25% with the lowest combined z score of the two measures ($N = 21$). The emotional *S*s made a mean of 21.8 CRs in the last 40 trials, while the nonemotional *S*s made a mean of 13.7 CRs. A t test gave a value of 2.18 which was significant beyond the .05 level. These results are consistent with the findings of the first study (Runquist & Ross, 1958). In a second analysis, the same pulse rate increase measure was used, but the skin conductance measure was the increase in base level conductance over the first 20 trials of conditioning. Samples of 18 emotional and 18 nonemotional *S*s were selected according to the same criteria as above.[2] The emotional *S*s made a mean of 24.5 CRs and the nonemotional *S*s a mean of 14.9 CRs in the last 40 conditioning trials. The t between these two means was 2.47 which was significant at better than the .05 level. These results confirm the findings of the second study (Runquist & Ross, 1959).

Muscle action potential indices.—Two measures of MAP activity were employed. The first was the base level of activity as determined by the sum of the MAP deflections in two randomly selected .5-sec. periods during the resting pretest period. Three groups of 20 *S*s each, representing high, middle, and low levels of such base level MAP measures were compared as to their conditioning performances. The mean number of CRs over the last 40 trials for these high, middle, and low groups were 22.9, 15.4, and 18.7, respectively. Analysis of variance of these data gave an F of 2.07, which fell short of significance.

The second measure of MAP related to level of conditioning performance was the response to the air puff. While the autonomic responses to the

[2] The smaller sample of emotional *S*s in this analysis resulted from the fact that some *S*s met the criterion on both measures. The nonemotional *S*s were then selected to keep the samples roughly the same size.

noxious stimulation have a latency of from 2 to 7 sec., it has been demonstrated (Davis, Buchwald, & Frankmann, 1955) that the latency of the MAP response is between .1 and .3 sec. A measure of MAP response to the air puff was taken as the summed deflections in a .5-sec. period following the air puff minus the summed deflections in a .5-sec. period preceding the CS. Nearly all *Ss* gave positive scores indicating an increase in MAP following the presentations of the air puff. By the tenth conditioning trial, however, the responses decrease to almost zero. For our present purposes, the sample of 86 *Ss* was divided into four groups corresponding to successive quartiles on the basis of the MAP response on the trial in which the air puff was given alone. The top and bottom quartiles contain 21 *Ss* per group; the two middle groups contain 22 *Ss*. Curves of acquisition of the CR in terms of percentage of CRs in each block of 10 conditioning trials are shown in Fig. 1. The curves clearly rank order themselves in terms of these MAP measures and show the divergence with trials implied by the theory concerning the interaction between *H* and *D*. Analysis of variance based on the total CRs for the last 40 trials gave an *F* of 8.99 which is significant beyond the .001 level. A rank order correlation between this MAP measure and number of CRs during Trials 41–80 gave a rho of .52.[3]

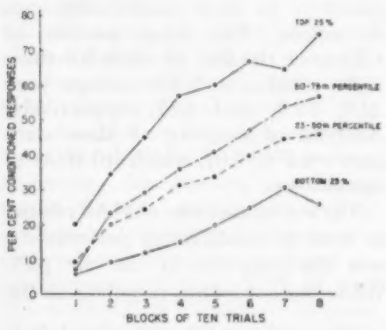

FIG. 1. Percentage of CRs in each block of 10 trials for four groups of *Ss* selected on the basis of the MAP response to the air puff alone.

DISCUSSION

The finding that emotionality as defined by autonomic response measures is related to performance in conditioning provides additional supporting evidence for the hypothesis that drive level (*D*) is a function of the level of emotional response of the individual. The second significant finding in this experiment is the high positive correlation between MAP response and performance in conditioning. Although the MAP measure may represent merely the magnitude of the skeletal startle response and not a change in any persisting tension level, this finding may, nevertheless, be interpreted as being consistent with the hypothesis that the four groups represent different drive levels. Thus, if it is assumed that the MAP response is some increasing function of the magnitude of this startle response to the air puff, and that the magnitude of the startle response is a function of drive level (*D*) then individual differences in MAP response should reflect individual differences in drive level. That the magnitude of the startle reflex is related to drive is, in fact, part of the theory of generalized drive, which states that any response will be facilitated by higher drive level. Experimental evidence for the effects of raising *D* on unconditioned reflexes has been provided by several studies (Brown, Kalish, & Farber, 1951; Meryman, 1953;

[3] The MAP measure could be combined with the autonomic measures to predict performance in conditioning. When this was done the 24 emotional *Ss* based on the 10 highest on each measure gave a mean of 23.9 CRs. The nonemotional *Ss*, the 22 *Ss* with the lowest total score, gave a mean of 14.6 CRs. The *t* between these means was 2.86, which is significant beyond the .01 level.

Spence & Runquist, 1958) which have shown increased unconditioned reflex amplitude when reflex occurs in the presence of a stimulus previously paired with some kind of noxious stimulus.

The writers who have used the concept of arousal level or activation in much the same manner as drive level has been used here, have postulated an inverted U relation between performance and arousal level, and have presented considerable evidence to support this generalization (Bartoshuk, 1955; Malmo, 1957; Malmo & Davis, 1956; Stennett, 1957). The results obtained here show that the conditioning performance is monotonically related to the magnitude of the MAP. However, the tasks used in the studies cited above have been relatively complex ones, presumably involving competing responses and other factors which are not analyzable according to the present theory. Such results are not contradictory to the theory and, considering that the conditioning situation is very different from those in which the U-shaped relationships have been obtained, the difference is not surprising. Furthermore, arousal level was defined in these studies as the increase in MAP occurring during a trial on these tasks, a definition which differs somewhat from the brief response to a noxious stimulus. There is no way to tell from the data of the present study whether these two MAP measures are related. It is felt, nevertheless, that the results are in agreement with the theoretical notions concerning the relation of drive level (D) to performance in eyelid conditioning and the hypothesis relating D to level of arousal as defined by MAP changes.

Summary

Eighty-six *Ss* were given 80 eyelid conditioning trials. During these trials pulse rate changes, skin conductance, and muscle action potential (MAP) were recorded. Two groups of *Ss* were selected on the bases of their GSR and pulse rate responses to the UCS. The group classified as emotionally responsive gave significantly more CRs than the group classified as nonemotional. Four groups selected on the basis of the magnitude of the MAP response to an air puff trial, were also compared. Mean number of CRs was found to be an increasing function of MAP response, the rank order correlation being .52.

The results were interpreted as supporting a theory that drive level in classical aversive conditioning is a function of the magnitude of the emotional response to a noxious UCS, and that the magnitude of MAP response is a direct function of drive level.

References

Bartoshuk, A. K. Electromyographic gradients as indicants of motivation. *Canad. J. Psychol.*, 1955, **9**, 215–230.

Brown, J. S., Kalish, H., & Farber, I. E. Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *J. exp. Psychol.*, 1951, **41**, 317–328.

Davis, R. C., Buchwald, A. M., & Frankmann, R. W. Autonomic and muscular responses and their relation to simple stimuli. *Psychol. Monog.*, 1955, **69**(20, Whole No. 405).

Duffy, E. H. The concept of energy mobilization. *Psychol. Rev.*, 1951, **58**, 30–40.

Freeman, G. L. The energetics of human behavior. Ithaca: Cornell Univer. Press, 1948.

Hebb, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.

Lacey, J. I. The evaluation of autonomic responses: Toward a general solution. *Ann. NY Acad. Sci.*, 1956, **57**, 123–164.

Lindsley, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 473–516.

Malmo, R. B. Anxiety and behavioral arousal. *Psychol. Rev.*, 1957, **64**, 276–287.

Malmo, R. B., & Davis, J. F. Physiological gradients as indicants of arousal in mirror tracing. *Canad. J. Psychol.*, 1956, **10**, 231–238.

Meryman, J. J. The magnitude of an unconditioned GSR as a function of fear conditioned at a long CS-UCS interval. Unpublished doctoral dissertation, State Univer. Iowa, 1953.

Runquist, W. N., & Ross, L. E. A preliminary study of the relationship between GSR, heart rate changes and level of performance in eyelid conditioning. *Proc. Iowa Acad. Sci.*, 1958, **65**, 398–404.

Runquist, W. N., & Ross, L. E. The relation between physiological measures of emotionality and performance in eyelid conditioning. *J. exp. Psychol.*, 1959, **57**, 329–332.

SCHLOSBERG, H. Three dimensions of emotion. *Psychol. Rev.*, 1954, **61**, 81–88.

SPENCE, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.*, 1953, **45**, 57–63.

SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.

SPENCE, K. W. An emotionally based theory of drive (D) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

SPENCE, K. W., & FARBER, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.

SPENCE, K. W., & RUNQUIST, W. N. The temporal relation of conditioned fear to the eyelid reflex. *J. exp. Psychol.*, 1958, **55**, 613–616.

SPENCE, K. W., & TAYLOR, J. A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

STENNETT, R. G. Relationship of performance level to level of arousal. *J. exp. Psychol.*, 1957, **54**, 54–61.

TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.

(Received January 30, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

INTERMITTENT PRIMARY REINFORCEMENT AS A PARAMETER OF SECONDARY REINFORCEMENT [1]

RICHARD M. KLEIN [2]

Boston University

A recent review of the literature on secondary reinforcement has clearly indicated the need for further empirical information of a basic kind (Myers, 1958). Little is known, for example, regarding the effects of schedules of reinforcement on secondary reinforcement, although a few studies suggest that there is some relationship (Saltzman, 1949; Zimmerman, 1957). Moreover, in view of the demonstrated significance of intermittent primary reinforcement in many learning situations, an investigator might expect that it would also be a parameter of the development of a secondary reinforcer. The present experiment was designed to investigate systematically one aspect of this possibility. More specifically, groups of *Ss* were given different frequencies of intermittent primary reinforcement during acquisition, a potential secondary reinforcer being present on every trial; then a test was made to determine the relative strength of the secondary reinforcer.

METHOD

Subjects.—The *Ss* were 72 male albino rats ranging in age from 90 to 120 days at the start of the experiment.

Apparatus.—The training apparatus was an alleyway, 45 in. long and 3½ in. wide on the inside. The sides and floor were made of wood and the top was covered with wire mesh. A guillotine door separated the starting box from the stem, and either a guillotine or a swinging door, depending on the stage of the experiment, separated the stem from the goal box. In length, the starting box was 9 in., the stem 24 in., and the goal box 12 in. The goal box was a detachable unit interlocked into the stem. Starting box, stem, and all doors were painted gray. Of the six goal boxes which could be fitted to the stem, two were painted black, two gray, and two white; and six glass food cups were painted the corresponding colors. The testing apparatus consisted of the training unit converted into a T maze by detaching the goal box, replacing it with a choice point chamber, and adding the original goal boxes to the chamber. The chamber was 6.5 in. long and 10 in. across. A T-shaped guillotine door was mounted in the chamber in such a way as to close off one arm of the maze from the other whenever it was down. Swinging doors were mounted in front of each goal box. Three 45-mg. food pellets served as reinforcements.

Training.—The *Ss* were fed 1 hr. daily for two weeks prior to the experimental runs. They were randomly divided into six groups of 12 each. All groups were given 20 training trials per day for six days, totaling 120, but they differed with respect to frequency of food-reinforced trials as follows: the 100% group received 20 per day; the 90% group, 18 per day; the 80% group, 16 per day; the 60% group, 12 per day; the 40% group, 8 per day; the 20% group, 4 per day. The preassigned schedules of reinforcement were entirely irregular during a day, from day to day, and from group to group. Each group was further divided into two groups of six each, one of which was trained on the black goal box and one on the white box. The day before training, *Ss* took a position preference test in the T maze with two gray goal boxes affixed. Three trials were given and two or more turns to a side counted as a preference for that side. On each training day a small group of *Ss* was brought into the experimental room and placed into individual

[1] This paper is based upon a dissertation submitted to the Graduate School of Boston University in partial fulfillment of the requirements for the Ph.D. degree. The author is deeply grateful to Garry Margolius whose constant aid and encouragement made this study possible.

[2] Now at United States Naval Training Device Center, Port Washington, N. Y.

intertrial cages. On the first reinforced trial *S* was permitted to stay in the goal box until the food was ingested. On all trials following, *S* was allowed only 10 sec. in the goal box. The mean intertrial interval was 5 min. The food cup was present on every trial. Prior to Trial 41 a guillotine door had been present in front of the goal box, but following Trial 41 it was replaced by a swinging door which was gradually lowered on each successive trial until (by Trial 48) it was fully vertical and was pushed up by *S*.

Testing.—The day after completion of training, *S*s were given 20 test trials. The T maze was assembled with white and black goal boxes. Each box contained its appropriate food cup, but food was never present during the test runs. For each *S* the goal box on which he was trained was placed against his position preference. The T-shaped guillotine door in front of the choice point and the swinging doors in front of each goal box prevented *S* from either viewing the goal box before entering it or from retracing. As was true during training, *S* was allowed 10 sec. in the goal box, and had an intertrial interval averaging 5 min.

Measures of secondary reinforcement.—Two measures were employed: the number of times *S* entered the training box as compared to the number of times it entered the novel box, and the percentage of times *S* entered the training box following first entrance into this box.

Results

The number of times each *S* of each group entered the training box during the test session is presented in Table 1, along with means and medians. The discrepancy between means and medians appears to be mainly due to the increase in extreme values, especially entries of 0 and 20, in the lower percentage groups. In the 40% group there are four values at the lower extreme and three at the upper extreme, and it is in this group that the mean and median diverge most sharply.

The Wallis-Kruskall *H* test was applied to the six groups, and the differences were significant ($P < .05$). A Mann-Whitney *U* test, two-tailed, was applied to every pair of groups.

TABLE 1

Number of Entries into the Training Box

*S*s in Each Group	Reinforcement Groups					
	100%	90%	80%	60%	40%	20%
1	14	16	19	20	20	20
2	14	15	19	20	20	20
3	13	14	18	19	20	20
4	12	14	17	19	19	20
5	12	13	15	19	19	20
6	12	13	15	16	19	19
7	11	12	14	16	17	18
8	11	11	14	14	16	15
9	10	10	12	12	0	14
10	9	10	11	10	0	12
11	8	10	11	10	0	10
12	8	6	9	0	0	0
Medium	11.5	12.5	14.5	16.0	18.0	18.5
Mean	11.2	12.0	14.5	14.6	12.5	15.7

Significant differences were obtained between the following pairs: 100% vs. 80% ($P = .02$), 60% ($P = .05$), and 20% ($P = .02$); 90% vs. 20% ($P = .05$). An examination of the data for individual *S*s reveals that three of the major groups, 60%, 40%, and 20%, include some *S*s who went either to the training box on every trial or to the novel box on every trial. But the first measure of secondary reinforcement was taken to be the number of times *S* entered the training box as compared to the number of times it entered the novel box. Hence the measure cannot be applied to the performance of the *S*s in question. They did not enter both boxes and a comparison of the reinforcing properties of the one as compared to the other cannot be made. A separate analysis was performed on the number of entries into the training box excluding all *S*s who went to the same box on every test trial. Consultation of Table 1 will reveal the number and place of the excluded *S*s. A Wallis-Kruskall test applied to the

six groups yielded significant differences ($P < .01$). Two-tailed U tests were performed on each pair of groups, and the following were found to be significant: 100% vs. 80% ($P = .02$), 60% ($P = .02$), 40% ($P = .002$), and 20% ($P = .05$); 90% vs. 40% ($P = .002$); 80% vs. 40% ($P = .05$).

Analysis of the data in terms of the percentage of times S entered the training box following the first such entry was performed. In this case too, and for the same reason given above, all Ss going to the same box on every test trial were excluded. However, before using this measure it must be reasonably assured that (*a*) the trial on which the training box was first entered occurred early in the test period for most Ss, and (*b*) the groups did not differ strikingly and systematically with respect to the trial number on which the first entry took place. Both of these requirements were fully met by the data. The first entry trials were all early; actually by Trial 5 almost all of the Ss included in this analysis had entered the training box at least once. The differences between groups were slight and no special trend was exhibited.

The mean and median percentages are shown in Table 2. The Wallis-Kruskall test gave significant differences ($P < .001$). The two-tailed U test yielded the following significant differences: 100% vs. 80% ($P = .02$), 60% ($P = .02$), and 40% ($P = .02$); 90% vs. 80% ($P = .05$), and 40% ($P = .002$); 80% vs. 40% ($P = .02$); 60% vs. 40% ($P = .02$).

TABLE 2

PERCENTAGE OF ENTRIES INTO THE TRAINING BOX FOLLOWING THE FIRST ENTRY INTO THE TRAINING BOX
(Does not include Ss who went to same box on every trial)

Measure	Reinforcement Groups					
	100%	90%	80%	60%	40%	20%
Median	60.0	62.5	81.0	92.0	100.0	96.0
Mean	61.1	63.9	76.7	85.5	99.2	84.5

The possibility that a greater number of Ss might have gone to the training box on Trial 1 in the lower percentage groups as compared to the higher percentage groups was analyzed, since, if this possibility materialized, the previous finding of a negative relationship between number of entries into the training box and amount of reinforcement might be confounded. A χ^2 test applied to the number of Ss going to the training box on the first test trial gave a probability between .50 and .30, and the null hypothesis cannot be rejected. Possible differences in test performance of Ss trained on the white goal box as compared to those trained on the black box were investigated. A two-tailed U test applied to each pair of brightness subgroups within each major group gave substantially high probabilities, and in no case could the null hypothesis be rejected.

DISCUSSION

The major finding of this experiment is that the pairing of a stimulus with primary reinforcement is a parameter of the secondary reinforcing properties of the stimulus when the stimulus is present on every trial and primary reinforcement occurs on only some of the trials. In general, it was found that decrease in the number of training trials on which primary reinforcement is present increases the strength of the secondary reinforcer. Some general characteristics of the function might be summarized. The 100% and 90% groups are not strikingly different from one another. Between the 90% and 80% reinforcement groups there is a rather abrupt increase in the percentage of entries into

the training box. From 80% to 20% there is an increase and leveling of the percentage of entries. On an empirical level, one can simply take the changes in the function as reflecting the strength of the secondary reinforcer. Also, current theories of intermittent and secondary reinforcement may be employed to explain the data in terms of theoretical concepts and perhaps also in terms of experimental variables other than, or in addition to, intermittent and secondary reinforcement. But it seems unlikely that any of the current theories would have been able to predict definitely the form of the obtained function. One obvious reason why this failure to predict may be true is that these theories are not sufficiently quantitative. In this regard, a promising formulation has been recently outlined by Wyckoff (1959). As applied to the present experiment, his theory would seem to imply that a major factor is the "cue strength" of the secondary reinforcer. He suggests that intermittent training would favor secondary reinforcement in that such training would not lead to a sudden change in the cue properties of the reinforcing stimulus at the commencement of extinction.

The data of this experiment revealed that some *Ss* went either to the training box or to the novel box on every test trial. Nothing similar to this phenomenon seems to be mentioned in the literature where a test similar to the one employed in the present experiment was utilized. Again, none of the current theories would seem to predict this phenomenon.

One of the more obvious implications of the present findings is with respect to the phenomenon of greater resistance to extinction shown by *Ss* who are intermittently reinforced as compared to those who are continuously reinforced. Our experimental results seem to suggest an empirical contingency which may help to explain this well known phenomenon. Stimuli present during intermittent reinforcement are apt to acquire more secondary reinforcing strength than stimuli present during continuous reinforcement. In the extinction test, the stimuli present would have stronger control over the behavior of the intermittently reinforced *Ss* than over behavior of the continuously reinforced ones. Hence resistance to extinction is greater in the former case. This explanation may be extended to cover even those cases where the stimuli are changed during extinction; for as long as the changed stimuli are on the same dimension as the training stimuli the possibility is present for generalized secondary reinforcement to gain some control over behavior. Presumably, these same hypotheses can also be applied to the finding that intermittently reinforced *Ss* tend to have poorer habit reversal than continuously reinforced Ss (Kendler & Lachman, 1958). The degree to which our findings are applicable to other studies is undoubtedly a function of many conditions, among which would be the extent to which these studies approximate our training and testing situations.

It is to be noted that the speculation given above is quite different from that first introduced by Denny (1956). His primary concern was with the effect of withdrawing secondary reinforcement during intermittent training, and he found that secondary reinforcement is a factor in an intermittent training situation. Our suggestion is that secondary reinforcement develops and generalizes during intermittent training and that its presence during extinction tends to prolong responding.

Summary

Six groups of *Ss* were trained to run an alleyway for intermittent reinforcement. The groups differed as to frequency of primary reinforcement, which ranged from 100% to 20%. On every training trial a potential secondary reinforcer, goal box brightness, was present. Following training, a test was made for the secondary reinforcing properties of the brightness stimulus by converting the alleyway to a T maze and attaching the training goal box and a novel one. In general, the strength of secondary reinforcement increased as number of reinforced

trials given during acquisition decreased. It was felt that the characteristics of the obtained function would not have been definitely predicted by any of the current theories, and the same appears to be true of the finding that some *Ss* went either to the training box or to the novel box on every test trial. The suggestion was made that the findings from this experiment may help to explain the frequent observation that resistance to extinction is greater following intermittent reinforcement than following continuous reinforcement.

References

DENNY, M. R. The role of secondary reinforcement in a partial reinforcement learning situation. *J. exp. Psychol.*, 1946, **36**, 373–398.

KENDLER, H. H., & LACHMAN, R. Habit reversal as a function of schedule of reinforcement and drive strength. *J. exp. Psychol.*, 1958, **55**, 584–591.

MYERS, J. L. Secondary reinforcement: A review of recent experimentation. *Psychol. Bull.*, 1958, **55**, 284–301.

SALTZMAN, I. J. Maze learning in the absence of primary reinforcement: A study of secondary reinforcement. *J. comp. physiol. Psychol.*, 1949, **42**, 161–173.

WYCKOFF, L. B. Toward a quantitative theory of secondary reinforcement. *Psychol. Rev.*, 1959, **66**, 68–78.

ZIMMERMAN, D. W. Durable secondary reinforcement: Method and theory. *Psychol. Rev.*, 1957, **64**, 373–383.

(Received November 28, 1958)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

THE EFFICIENCY OF UTILIZATION OF VISUAL INFORMATION AND THE EFFECTS OF STRESS [1]

AUSTIN JONES

University of Pittsburgh

Attneave (1954) has suggested an information model of visual perception which is the conceptual basis for the research reported here. In Attneave's analysis, the visual field may be regarded as constituted by very many small cells. There are several classes of cells, corresponding to the visual "ground" and to each of the one or more objects which comprise the "figure." As the organism scans the visual field, implicit probability operations are performed upon the status of cells. The probability of a cell falling in a particular class is conditional upon the status of adjacent cells; if, for example, a particular cell falls within a certain object's contour, the probability is increased that adjacent cells also fall within the contour. In this way it becomes possible to demonstrate redundancy within the field; the organism need not scan or "see" all the cells in order to predict or "perceive" the particular object which is represented there.

Stated in its most general form, the purpose of the present study is to assess the effects of stress upon the efficiency of utilization of visual information. "Visual information" refers to the potential that cells contain for the reduction of uncertainty in the organism as to which of a population of contours is represented in the field. Efficiency of utilization of visual information is inversely related to the proportion of the total cells in a contour which must be presented to the organism in order for a criterion perception to be attained. An organism correctly "recognizing" a particular contour on the basis of very few of the total possible cells is thus regarded as more efficiently utilizing visual information than an organism requiring a greater number of cells. In order to measure such efficiency of utilization of visual information, a new technique was developed involving the use of what will be referred to as *probability grid figures*.

The study most closely related to the present one is that reported by Eriksen and Wechsler (1955) on the effects of anxiety upon visual size discrimination as measured by the method of absolute judgment. Their *S*s were asked to designate, by rank, the size of 11 different squares as they were presented in random order for brief intervals. A prior training period consisted of two serial presentations of the squares during which *E* called out the correct rank of each square. The principle dependent measure was I_T, information transmitted, which expresses the discrimination accuracy of *S*s' responses. Anxiety was generated by the administration of strong electric shock at random intervals. No differences were obtained in discrimination accuracy between anxious *S*s and nonshocked controls. "Both groups transmitted 1.87 bits of information about the stimulus series, indicating that they could absolutely discriminate slightly less than four stimuli within the size range investigated" (Ericksen & Wechsler, 1955, p. 462). It was found, however, that the anxious group was "more

[1] This report is based upon a thesis submitted to the Graduate School of the University of Rochester in partial fulfillment of the requirements for the Ph.D. degree. The author is indebted to his advisor, Emory L. Cowen, and to Robert Boynton, Russel Green, and Morton Wiener for their helpful criticism and advice.

biased in the use of the 11 response numbers and at the same time . . . more predictable or stereotyped in the responses they applied to specific stimuli" (Ericksen & Wechsler, 1955, p. 462).

The purpose of the present study is to relate efficiency of utilization of visual information to variations in intensity of stress as induced by electric shock. Four levels of shock were employed, ranging from a "minimum perceptible" level to the "maximum tolerable" level as established for each *S*. The hypothesis tested was that efficiency of utilization of visual information is a decreasing monotonic function of stress. An additional purpose of the study was the methodological one of demonstrating the utility of the *probability grid figures* for the study of visual information processes.

METHOD

Probability grid figures.—The dependent variable of the study is a score representing each *S*'s response to the probability grid figures. The development of the figures is described in detail elsewhere (Jones, 1957). A brief resume of the method is given here.

The materials consist of eight familiar visual forms, each of which is presented in a series of separate stimulus figures graded according to the proportion of total area of the form that is reproduced upon them. This was accomplished by constructing the form of black cells on a ground of white cells and randomly drawing successive percentages of the total black cells to be reproduced as the successive figures of the graded series. The randomness here refers to the location within the total form of the black cells selected. The scale includes such forms as the letter H, an arrow, the Arabic number 88, etc. In Fig. 1 are shown some examples of particular forms at different percentage-of-information levels.

All *S*s are presented with all stimulus figures in a prearranged fashion. First are presented all figures of the lowest level of information, i.e., of the lowest proportion of total area reproduced. This includes all eight forms in a randomized order. Then are presented all figures of the next higher

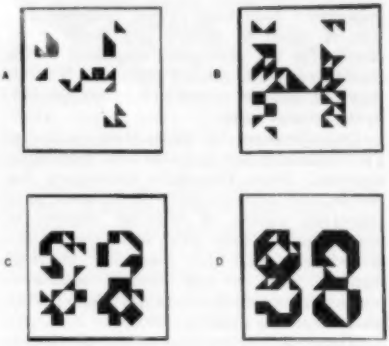

FIG. 1. Examples of probability grid figures: (*a*) the letter H reproduced at 20% of the total information; (*b*) the letter H reproduced at 40% of the total information; (*c*) the number 88 reproduced at 55% of the total information; (*d*) the number 88 reproduced at 75% of the total information.

percentage of information, with the order of forms randomized anew. The procedure continues in this fashion until all figures of the uppermost level (percentage) of information have been presented. The *S* is required, upon the presentation of each figure for a maximum of 10 sec., to name or describe what he has seen. He is told that the incomplete figures will become clearer as he goes along but that he must respond whether or not he is sure of his answer. The value of the lowest percentage level at which *S* correctly perceives a given form constitutes his "efficiency of information utilization score" for that form; for further ease of discussion, this will be referred to simply as the "information score." The information score for the total scale is the mean score for the eight forms. The function of *S*'s verbal response is simply to serve as a sign to *E* that the form has been perceived.

The exact instructions, as given to the *S*s of this study, presented the task as involving "very simple patterns or designs . . . that are very familiar to everybody," and discouraged *S*s from regarding the procedure as a test of imagination or personality in any way. From Fig. 1 it may be seen that there is a similarity between the probability grid figures and such earlier procedures as the Street Gestalt Test (Street, 1931), which employed fragmented pictures of common objects. The Street figures were not constructed so as to yield formal information

measures. Thurstone (1944) proposed factors of speed of closure and flexibility of closure for the perceptual responses to the Street figures and related tasks, and it seems that such concepts would be equally applicable in the present case.

Quantification in information terms of the visual stimulus has also been previously reported. Fitts, Weinstein, Rappaport, Anderson, and Leonard (1956), for instance, have generated figures of differing degrees of redundancy. These differ from the present procedure in that they are solid, not fragmented, and do not involve sequential increments of information regarding a particular, commonly observed form.

Apparatus.—The expectation of electric shock, reinforced at irregular intervals, constituted the stress variable of the study. Shock was administered through ordinary silver electrodes strapped to *S*'s forearm. Voltage was regulated by a Variac. A Hunter timer controlled the duration of shock.

Subjects.—The *S*s were 37 male undergraduates at the University of Rochester. They were assigned in rotated order to four groups corresponding to the four shock (stress) levels to be used—zero shock, minimum preceptible shock, pain midpoint shock, and maximum tolerable shock—hereafter referred to, respectively, as Groups 0, 1, 2, and 3. Ten *S*s fell in Group 3, nine in each of the others.

The *S*s were told that the experiment related to the development of a new test of visual perception. They were told also that unpleasant electric shock was a part of the procedure but that the maximum shock they could receive would be set by themselves beforehand. The *S*s assigned to Group 0 were told prior to the experimental session that they had been selected to serve as control *S*s and would not receive any shocks throughout the procedure.

Procedure.—All *S*s were seen individually by *E*. Group 0 was administered the probability grid figures in the same experimental room as were the others. For Groups 1, 2, and 3, the procedure was as follows. The electrodes were placed on *S*'s forearm and a single ascending series of shocks administered, beginning with the Variac set at 4 v., and proceeding at 4-v. intervals. The *S* was instructed to tell *E* when he could first feel any sensation and when the shock became intolerably strong. The level of shock which *S*s of Groups 1, 2, and 3 received in the subsequent experimental procedure was determined in this way: For Group 1, the Variac was set at the lowest level at which any sensation was reported. This was typically described by *S* as a faint tickling, not in any sense painful. For Group 3, the Variac was set at the highest level prior to that judged to be intolerable. The *S*s of Group 2 were given a level of shock intended to approximate the midpoint of pain between these extremes. Prior pilot work related to pain scaling indicated, for the particular apparatus involved, that the subjective midpoint in pain occurs rather reliably at a voltage that is equal to the threshold voltage plus one-third the voltage difference between threshold and the level denoted as maximum tolerable. In the present study the "pain midpoint" for *S*s of Group 2 was set according to that formula.

After establishing *S*'s maximum tolerable shock, he was told that in the experiment to follow he would receive several shocks which would all be of the same intensity. He was then given a single "sample" shock of the intensity he would later receive, the intensity being set according to the experimental group (1, 2, or 3) to which he was assigned.

The probability grid figures were then administered. As in the case of Group 0, *S*s sat across a desk from *E*, viewing the figures from a distance of 3 ft. Ten electric shocks of 1.5-sec. duration were interpolated at irregular intervals into the course of *S*'s responses to the 136 figures. The *S*s understood that the procedure did not constitute a timed test and that their time did not affect in any way the shocks they received. Following the presentation of the last probability grid figure, *S*s were interviewed briefly to resolve any negative feelings toward the experiment and to secure their help in maintaining the confidentiality of the procedure.

Results and Discussion

Figure 2 presents graphically the effects of the several stress conditions upon efficiency of utilization of visual information. The data were analyzed by an analysis of variance procedure of the sort described by Grant (1956) for the evaluation of trends. The analysis, summarized in Table 1, shows a significant F ratio for the linear component of the between groups variation ($F = 8.66$,

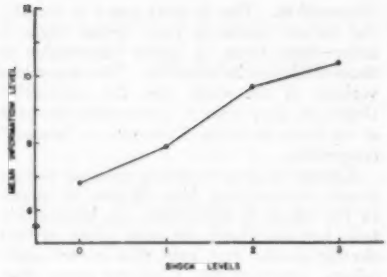

FIG. 2. Efficiency of utilization of visual information as a function of stress.

$P < .01$). The hypothesis of the study was supported; efficiency of utilization of visual information was shown to be a decreasing monotonic function of stress. High information scores, it will be remembered, are indicative essentially of high recognition thresholds and represent less efficient utilization of information. The relationship between information scores and stress, as graphed in Fig. 2, appears to be essentially linear. In the analysis of variance, the linear component accounts for almost all of the variance between groups. A product-moment correlation was computed between information scores and stress levels; r (significant at $<.01$) = .46.

The results of the present study and those of the study by Eriksen and Wechsler (1955) cited earlier are in interesting contrast. In the Eriksen and Wechsler study both the anxious and the nonshocked control groups transmitted almost exactly the same amount of information about the stimulus series—1.87 bits. In the present study, impairment of information processing seems clearly demonstrated, and inspection of the total function in Fig. 2 suggests that even relatively minor degrees of stress, as were associated with Group 1, may have a significant decremental effect. The two studies are roughly comparable in terms of their principal variables. The dependent variable, the information score of the present study, appears to be essentially parallel to the dependent variable I_T, information transmitted, of the earlier study. Each expresses the extent to which information about a stimulus series is transmitted through the human channel.

The independent variable, stress, was, of course, arranged quite similarly in each study by the use of electric shock. The differences in experimental results, then, must reflect differences in the attributes of the two information tasks involved, attributes which remain to be specified by further study.

It is not within the scope of the present study to evaluate the characteristics of *S*s' responses associated with decreased efficiency of information utilization. Some informal observations along this line were made, however. Certain *S*s, particularly in Group 3, produced a striking proportion of perseverative, incorrect responses. In the most extreme case, for instance, *S* responded "square" uniformly and incorrectly to four of the eight forms up through their seventeenth and final presentation. Figure 3 shows the mean frequency with which the five most frequently repeated incorrect responses were given for Groups 0 and 3. The *S*s of Group 3, it will be noted, responded, on the average, 23 times with their most frequently repeated incorrect response (regardless of its content), while the *S*s of Group 0 responded with their most frequently repeated incorrect responses (regardless of its content) only 12 times. For the next four most fre-

TABLE 1

ANALYSIS OF VARIANCE OF VISUAL INFORMATION SCORES

Source	*df*	*MS*	*F*
Between groups	3	25.27	2.65
linear	1	73.33	7.69*
residual	2	1.20	
Within groups	32	9.53	
Total	35		

* $P = .01$.

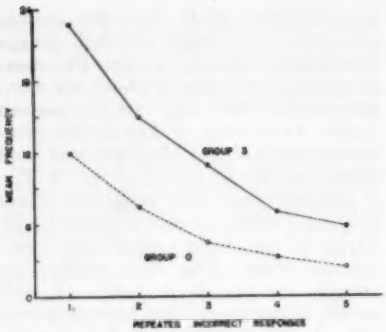

FIG. 3. The mean frequency with which the five most frequently repeated incorrect responses were given by the *Ss* of Groups 0 and 3.

quently used incorrect responses, a similar relationship is maintained. Such stereotyping of response corresponds to Eriksen and Wechsler's (1955) observation of the lowered response information I_R of their shocked *Ss* relative to controls, i.e., their tendency to bias in the use of the response categories that were assigned for the task.

A further concern of this study has been the demonstration of the potential utility of the probability grid figures used here for the first time. In view of their apparently meaningful use in the present study, it is suggested that the information score associated with response to the figures might be advantageously employed as the dependent variable in a variety of other studies having to do with information processing in visual perception.

SUMMARY

This study was designed to test the hypothesis that psychological stress reduces the efficiency of information processing in visual perception. Toward this end, a new information scaling procedure was devised. A variety of simple forms were each reproduced in a series of 17 ascending levels of information. The *Ss* were asked to identify the various forms as they viewed them in order—from those of lowest information to those of highest information. The dependent variable of the study was the recognition threshold, that level of information for each of the forms at which *S* first gave a "correct" recognition.

College *Ss* were randomly assigned to four groups representing four degrees of stress. In the course of performing the information task just described, *Ss* were given several electric shocks and told that more might follow. Group 0 was a control group that received no shock; Groups 1, 2, and 3 received minimum perceptible, moderate, and painful shocks, respectively. The principle analysis was a comparison of the mean thresholds for each group, i.e., the levels of information at which correct form recognition was first achieved. The results showed significant differences ($P < .01$) among the four groups, with increasing degrees of stress resulting in corresponding increases in recognition thresholds—or decreases in the efficiency of utilization of visual information.

REFERENCES

ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193.

ERIKSEN, C. W., & WECHSLER, H. Some effects of experimentally induced anxiety upon discrimination. *J. abnorm. soc. Psychol.*, 1955, **51**, 458–463.

FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, N., & LEONARD, J. A. Stimulus correlates of visual pattern recognition: A probability approach. *J. exp. Psychol.*, 1956, **51**, 1–11.

GRANT, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

JONES, A. The efficiency of utilization of visual information and the effects of stress. Unpublished doctoral dissertation, Univer. of Rochester, 1957.

STREET, R. F. A Gestalt completion test. *Teach. Coll. Contr. Educ.*, 1931, 481.

THURSTONE, L. L. *A factorial study of perception.* Chicago: Univer. Chicago Press, 1944.

(Received December 4, 1958)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

THE PERCEPTION OF REPETITION RATE[1]

JOSEPH C. STEVENS AND GERARD M. SHICKMAN

Harvard University

This research took its departure from two experiments, one by Reese (1943) on the perception of a flashing light, and the other by Pollack (1952) on the perception of a regularly interrupted noise. Both these experiments sought to determine how the *apparent* or *subjective* frequency of a regularly repeated signal grows as a function of its *actual* rate of repetition. On the basis of the fractionation data obtained in these experiments, S. S. Stevens (1957b) observed that apparent frequency grows approximately as the 1.7 power of the repetition rate. The present experiments confirm the existence of the power function, but they show that the exponent is not appreciably different from unity.

This first-power law governing apparent frequency seems to be essentially independent of the modality stimulated, at least for light flashes, noise bursts, clicks, repetitive electrical pulses and repetitive mechanical pulses to the fingers. Nor is it greatly influenced by signal intensity or by the relation between pulse duration and period (i.e., duty cycle).

These problems were investigated by the method of fractionation (halving and doubling) and by the method of magnitude estimation. In addition, a category-rating scale was obtained for each modality in order to determine whether a plot of category judgments against ratio judgments of repetition rate would exhibit the same convexity as do such plots for brightness, loudness, and other prothetic continua. A similar comparison was also made between the JND scale and the ratio scale.

[1] This research was supported partly by Grant G-2668 from the National Science Foundation and partly by Contract Nonr-1866(15) with the Office of Naval Research (Project Nr142-201, Report PNR-222). Reproduction for any purpose of the United States Government is permitted.

METHOD

Apparatus

A square-wave generator determined the repetition rates of visual, acoustical, electrical, or mechanical pulses presented to the observer (*O*) through auxiliary equipment. The pulses themselves were delivered to *O* by an appropriate "stimulator": salt water electrodes (for shocks to the fingers), a vibration generator, a pair of earphones, or a flash source. The "stimulator" could be driven directly by the square-wave generator (for pulses of constant duty cycle) or by a pulse generator triggered by the square-wave generator (for pulses of constant duration). The *O* could initiate stimulation by pressing a switch that operated a timer in the driving circuit. Figure 1A illustrates the relations among these components.

In one phase of the experiment, the output of a white-noise generator was interrupted at frequencies determined by the internal multivibrator of an electronic switch and presented to *O*s through earphones.

For experiments involving halving and doubling, the apparatus was altered so that *O* could control the frequency of stimulation by adjusting a continuously variable audio oscillator. A decimal scaling unit emitted one square pulse for every 100 cycles of the oscillator output. These pulses and the standard frequency from the square-wave generator were switched alternately into the pulse generator by a motor-driven commutator, and the resulting pulse trains, applied to the appropriate "stimulator," gave rise to corresponding standard and variable sequences of light flashes or mechanical pulses. To produce interrupted white noise having a constant duty cycle, the output of the pulse generator was converted into series of square waves that gated an electronic switch. The signal from a white-

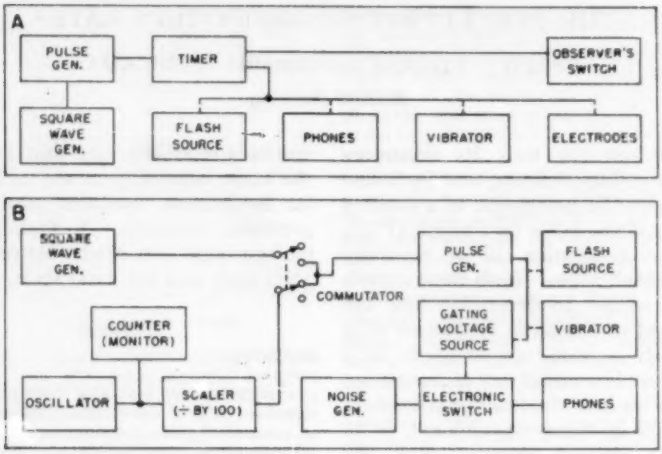

FIG. 1. Generalized block diagram of apparatus for (A) magnitude estimation and category scaling and (B) fractionation. Only one of the stimulators shown was included at any time. Interchangeability is indicated by dashed lines.

noise generator was passed through the switch and transmitted to *O* via PDR-8 earphones. The *O*'s settings of variable stimuli were monitored continuously by a decimal counting unit at the output of the oscillator. This apparatus is illustrated schematically in Fig. 1B.

Visual Flash Rate

A Sylvania R-1131C glow-modulator tube was shielded from *O* by a black partition so that only a circular flashing target, subtending ½ degree of arc, was visible. The *O* looked at the target through a binocular aperture at one end of a black hood, about 35 cm. long. To *O* each train of stimuli appeared as a series of white flashes against a dark background.

The duration of the square pulses to the glow-modulator tube was 300 μsec. throughout all the experiments. At a flash frequency of 60/sec., the photometric level of the target was about 1 lambert.

Magnitude estimation.—Twelve *O*s, students and employees of the laboratory, participated in two separate sessions. The same set of stimuli was presented in both sessions, but for the second session the photometric level was reduced a thousandfold by a neutral density filter.

The *O* was given written instructions to let the number 10 stand for the apparent frequency of the first stimulus and to assign numbers to subsequent stimuli in proportion to their apparent frequency. Seven stimuli between 1 and 40 flashes/sec. were presented twice each in different irregular orders to each *O*. The flash rate of the first stimulus, called 10, was 9/sec. By depressing a switch *O* initiated the stimulus, whose duration was limited to about 3 sec.

Halving and doubling.—Five of 10 *O*s made judgments of half-frequency in the first of two sessions, and of double frequency a day or two later; the other *O*s followed the reverse order. Standard and variable stimuli alternated in a repeating cycle (standard, 4 sec.; darkness, 1 sec.; variable, 4 sec.; darkness, 1 sec., etc.). The instructions were to adjust the variable stimulus to a frequency that appeared to be one-half or double that of the standard. The *O* was requested to approach the final setting by bracketing, i.e., by turning the knob of the oscillator successively too high and too low. He was further requested to refrain from counting the flashes when counting was possible. The five standard rates presented ranged from 2 to 25 flashes/sec. (halving) and from 1 to 16 flashes/sec. (doubling). The standards appeared twice each in irregular order.

Category rating.—Ten *O*s made category judgments of the same set of stimuli used in the estimation studies. The instructions were to assign each stimulus to one of seven

categories, designated in order of increasing frequency by the numbers 1 to 7. At the beginning of the session the slowest (1 flash/sec.) and the fastest (40 flashes/sec.) stimuli were presented and identified for *O*s as the extremes of the stimulus range. Again, *O* initiated stimulation, which lasted about 3 sec. Each *O* received the stimuli twice in different irregular orders.

Auditory Flutter Rate

In one estimation experiment, the stimulus consisted of fairly loud, discrete clicks, presented binaurally through PDR-8 earphones excited by 16-μsec. positive pulses. Ten *O*s estimated the apparent frequency of seven repetition rates between 1/sec. and 40/sec. The repetition rate of the first stimulus, 9/sec., was called 10.

In a second estimation study, the stimulus was interrupted white noise, whose on-off ratio remained constant at 0.5 regardless of frequency. The amplified output of an electronic switch passed through an attenuator that determined the signal level in PDR-8 earphones at 100 db above 0.0002 microbar (for one part of the experiment) and at 50 db (for another part). Twelve *O*s reported the apparent frequency of 10 stimuli between 1/sec. and 60/sec. Half of the *O*s listened first to the signals of high intensity and later to those of low intensity, and the other half listened in the reverse order. The stimuli were presented in two different irregular orders to each *O*. The rate of the first stimulus, called 10, was 10/sec.

For halving and doubling, the stimulus was interrupted white noise with a constant on-off ratio of 0.5, presented binaurally at 85 db above 0.0002 microbar. Five *O*s made judgments of half-frequency in the first of two sessions and of double frequency in the second, and five followed the reverse order. The five standard rates ranged from 2/sec. to 64/sec. (halving), and from 1/sec. to 32/sec. (doubling).

In addition, eight *O*s made category judgments, using a 7-category scale, of white-noise pulses at 100 db sound pressure level. The spacing of the stimuli was roughly geometric between 1/sec. and 40/sec. To test the effect of stimulus spacing on the form of the category scale, eight additional *O*s made category judgments of a set of stimuli spaced at equal intervals of repetition rate.

Tactual Pulse Rate

The *O* rested his right middle finger tip lightly upon a plastic button about ¼ in. in diameter affixed to the armature of a vibration generator. The vibrator was actuated by positive square pulses, whose duration was fixed by the pulse generator at 1.7 msec., and whose amplitude was adjusted to produce a pulsating sensation of moderate intensity.

As in the preceding experiments, *O* pressed a switch that triggered the stimulus for about 3 sec. (in magnitude estimation and category rating), or received in alternation standard and variable stimuli whose presentation was timed by the motor-driven commutator (in halving and doubling).

Eleven *O*s made magnitude estimates, eight *O*s made category rating judgments on a 7-point scale, and ten *O*s made fractionation judgments (five *O*s beginning by halving, the other five by doubling). All but a few of the *O*s had taken part in one or more of the other experiments on apparent frequency.

The instructions, stimulus rates, and procedure were similar to those described in the section on flash rate. In magnitude estimation, repetition rates of the mechanical pulses ranged in seven steps from 1/sec. to 40/sec., the stimulus called 10 being 9/sec. In fractionation, the five standard rates ranged from 1/sec. to 20/sec. for doubling, and from 2/sec. to 40/sec. for halving. In category rating, the same pulse frequencies were used as in magnitude estimation.

Electrical Pulse Rate

Square waves of constant duty cycle, or electrical pulses of constant duration, were impressed across the first and third fingers, which were immersed up to the first joint in salt solutions energized through submerged aluminum plates. A resistance of 10,000 ohms was placed in series with the pulse generator and *O*, and a resistance of 75,000 ohms in series with the square-wave generator and *O*. The peak stimulating potential, as measured across a 22,000-ohm resistor substituted for *O*'s fingers, amounted to 15 v. People vary, of course, with respect to skin resistance, but under these conditions the typical *O* reported a moderately intense, clearly pulsating sensation. (Fingers having cuts or abrasions were not used.) In all experiments, *O* pressed a switch to initiate the stimulation, which persisted about 3 sec.

In the first experiment 18 *O*s estimated the apparent frequency of nine stimuli whose on-off ratio was constant at 0.5, and whose actual repetition rate ranged from 2/sec. to 30/sec. The first stimulus, called 10, was 12/sec.

Seven of the same *O*s and three others took part in a second experiment, this time estimat-

ing the frequency of stimuli whose duration remained constant at 1.7 msec. Seven repetition rates from 2/sec. to 60/sec. were used; the rate of the first stimulus (called 10) was 9/sec.

In a third experiment, eight *Os* made category judgments, on a 7-point scale, of the set of stimuli used in the second experiment.

Results

Magnitude Estimation

The results of the estimation studies appear as medians in the log-log coordinates of Fig. 2. The size of the interquartile ranges, as represented by the vertical lines, indicates that apparent rate was not estimated with high precision. Despite the variability, however, the general form of the functions is clear. The linear alignment of the medians in log-log coordinates shows that apparent frequency grows as a power function of actual repetition rate, or

$$F = kR^n$$

where F is apparent frequency and R, repetition rate. The empirical ex-

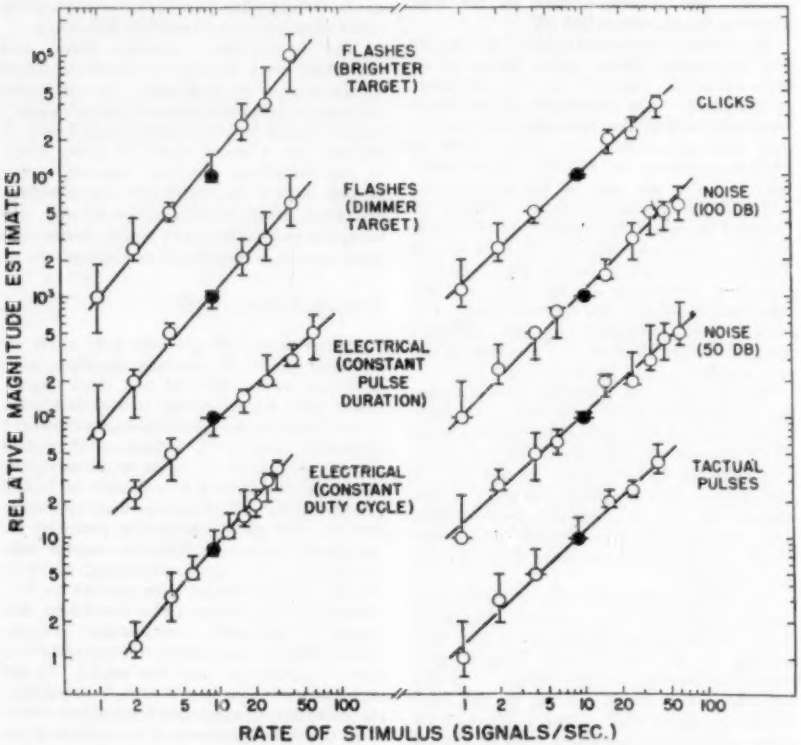

FIG. 2. Magnitude estimates of apparent frequency. Medians and interquartile ranges; each point represents two judgments by each *O*. Filled circles indicate estimates of stimulus rates originally identified as "10." In order to separate the lines along the ordinate the *y* values have been multiplied (from bottom to top line in each column) by 1, 10, 100, and 1000, respectively.

ponent n, as estimated from the slopes of the lines in Fig. 2, is close to unity for flashes, clicks, interrupted noise, and tactual pulses.

The intensity level of the signal does not appear to affect the growth of auditory flutter rate. Neither does it make much difference to the form or slope of the flash-rate function, although at low intensity the two highest rates appeared to a few *O*s as fused or barely flickering. (If *O* could not make an estimate, or estimated the apparent frequency as zero or infinite, his judgments of these stimuli were disregarded in taking medians.)

The difference between stimuli of constant pulse duration and stimuli having a constant on-off ratio does not affect the slopes of the auditory functions, although it may influence slightly the functions for electrical pulse rate. The slope (exponent) of the function governing electrical stimuli of constant on-off ratio turns out to be about 1.2, whereas that of the function for stimuli of constant pulse duration is approximately 0.9. Like those of the other rate functions, these slopes differ little from unity. It would be interesting to know whether the slight difference in slope, corresponding to the two types of electrical stimulus, would stand up under replication. The difference does not appear to relate to the number or identity of *O*s, since the same difference in slope turned up in the estimates of the seven *O*s who judged both kinds of electrical stimuli. On the other hand, since the electrical stimuli differed in number, range, and spacing in the two experiments, it is perhaps hasty to attribute the difference of the slopes to the temporal patterns of stimulation. The auditory functions are alike, despite such variations in temporal pattern. In any event, the fact that the slopes of all the functions are close to unity signifies that the apparent frequency of electrical pulses, like the apparent frequency of pulsating auditory, visual, and tactual stimuli, grows roughly in proportion to actual repetition rate.

Halving and Doubling

Figure 3 combines the median results of the halving and doubling procedures. The ordinate is the rate

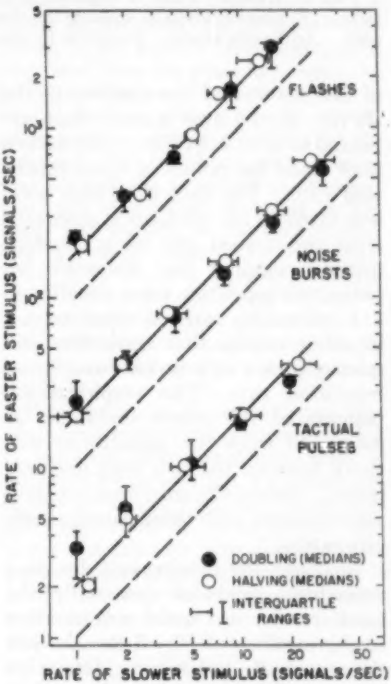

FIG. 3. Halving and doubling of apparent frequency. *Open circles*: medians of half-judgments. *Filled circles*: medians of double judgments. *Vertical and horizontal segments*: interquartile ranges. Solid lines fit by eye to composite halving and doubling data. Dashed lines show equal frequency. In order to separate the lines along the ordinate, the y values have been multiplied (from bottom to top) by 1, 10, and 100, respectively.

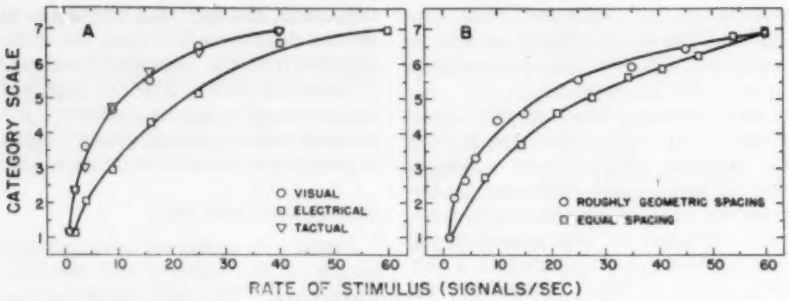

FIG. 4. Category scales of apparent frequency: (A) flashes, electrical pulses, and tactual pulses; (B) interrupted noise, showing the effect of stimulus spacing on the form of the category scale. Arithmetic means. Curves fit by eye.

of the faster, and the abscissa of the slower, of the two stimuli that appeared to be in 2:1 ratio. The dashed lines stand for points of equal repetition rate. The fact that each solid line fitted to the medians is separated from the dashed line by a constant distance signifies that the ratio between two repetition rates standing in 2:1 subjective ratio is constant, or, in other words, that subjective frequency grows as a power function of repetition rate. The empirical exponents of the power functions, as estimated from the logarithmic distance between the two lines, are near unity. Hence the fractionation data are consistent with those of magnitude estimation.

Although the halving and doubling procedures produced essentially the same result, a careful examination of the medians of Fig. 3 reveals that the one-half judgments determine lines of slightly steeper slope than do the judgments of double apparent frequency. In other words, *O* tended on the average to err in the direction of setting the variable stimulus toward intermediate rate values. The direction of this discrepancy coincides with similar discrepancies found in fractionation judgments of other sensory magnitudes (S. S. Stevens, 1957a), and confirms a recommendation (S. S. Stevens & Poulton, 1956) that in fractionation procedures, halving should be balanced by doubling. By fitting the lines in Fig. 3 to points for halving and doubling collectively, we strike an average between the results of the two procedures and effectively compensate for *O*'s tendency to "regress" to moderate levels.

This "regressive" tendency is more pronounced in adjustments of flutter rate than those of flash rate and is still more extensive with tactual pulses. Most of the *O*s who judged tactual pulse rate complained about the difficulty of their task: often in the course of a prolonged session, the discrete character of the pulses seemed to fade, and it became necessary for *O* to rest or for *E* to increase the amplitude of the pulses. Although amplitude as such does not appear to be an important determinant of the function relating apparent to actual repetition rate, it is quite conceivable that the precision of *O*'s judgments was affected by distractions arising out of these variations in the effectiveness of stimulation.

Category Rating

Figure 4 shows mean category assignments (ordinate) as a function of repetition rate. Inasmuch as the results of magnitude estimation and fractionation show that apparent rate closely follows actual repetition rate, the form of the functions would not be appreciably altered if the category assignments were plotted against apparent rather then actual frequency. Like category functions for such prothetic continua as loudness and brightness (S. S. Stevens & Galanter, 1957), the curves for the repetition rates of flashes, clicks, bursts of noise, tactual pulses, and electrical pulses all exhibit a convex shape, indicating that the category rating procedure fails to yield a scale of equal intervals of apparent frequency.

As Fig. 4B shows, the way in which the stimuli are spaced affects the form of the category scale. Spacing by equal intervals of repetition rate resulted in a slightly less convex function for interrupted noise than did spacing by roughly equivalent geometric steps. A similar dependence of the category scale upon stimulus spacing has been noted for many perceptual continua (S. S. Stevens & Galanter, 1957) and contrasts with the near invariance of magnitude estimates under radically different kinds of stimulus spacing (J. C. Stevens, 1958).

Discussion

The findings of this study support three major conclusions: (*a*) apparent frequency is related to physical repetition rate by a power law with a smaller exponent than was suggested by earlier investigations; (*b*) apparent frequency displays characteristics of a prothetic, rather than a metathetic, psychological continuum; and (*c*) ratio scales of apparent frequency obtained with different kinds of stimuli are markedly consistent.

As already noted, the results of the studies by Reese on apparent flash rate and by Pollack on apparent flutter rate do not agree with the conclusion that apparent frequency grows in proportion to repetition rate. Although the fractionation data from both these earlier experiments can be described fairly well by power functions, the exponents are definitely larger than those obtained in the present studies. Whether this disagreement is attributable to the small number of *O*s (five) employed in the earlier experiments, or to some other factor, is not known. One such factor may have been the procedure used. In both experiments *O* was required to choose, from an array of "fixed" stimuli, one whose rate appeared half that of the standard. It has been well demonstrated that the judgment of apparent ratio can be seriously distorted by the range and spacing of the fixed stimuli that are made available to *O* (Garner, 1954; S. S. Stevens, Rogers, & Herrnstein, 1955). By providing *O* with a continuously variable control over a very wide range of stimulus magnitude, we presumably avoid this particular source of bias.

Both the fractionation and estimation procedures used in the present study led to essentially the same ratio scale of apparent frequency, but the scales obtained by category rating differed significantly from it. Nonlinearity between the category scale and the ratio scale of apparent magnitude is typical of prothetic continua, a class to which apparent frequency seems to belong. The reasons for this nonlinearity are discussed by S. S. Stevens and Galanter (1957).

The prothetic character of apparent frequency is also evidenced by the fact that the ratio scale does not seem to be proportional to the scale obtained by the summation of JNDs. JND scales were constructed by Gebhard, Mowbray, and Byham (1955, Table 1, p. 53), and

Mowbray, Gebhard, and Byham (1956, Table 1, p. 107), who used the method of adjustment to measure differential sensitivity (in terms of the average deviation) for the repetition rate of flashes and noise bursts. The summated JND scales thus obtained show little resemblance to the power function, $F = R^{1.0}$, which represents the generalized outcome of the scaling of apparent frequency by magnitude estimation and fractionation. As a matter of fact, the forms of the summated JND scales for flash and flutter seem to differ from each other, as well as from the ratio scale. In contrast, the form of the ratio scale is relatively independent of the nature of the stimulation.

Summary

The apparent frequency of periodic flashes of light, bursts of noise, clicks, tactual pulses to the finger tip, and electrical pulses applied across the fingers has been found, by the methods of fractionation and magnitude estimation, to grow approximately as the first power of physical repetition rate, at least within the range from 1/sec. to 60/sec. This relation is virtually unaffected by the intensity of stimulation, and it holds whether the duration of the pulses is independent of, or directly proportional to, their period.

Apparent frequency displays properties of a prothetic psychological continuum, in that scales of apparent frequency obtained by category rating and by summation of JNDs are both nonlinear, relative to the ratio scale determined in this study. The form of the ratio scale, unlike that of JND scales of apparent frequency, is not materially dependent upon the modality stimulated and, unlike the form of category scales, is not influenced by the spacing of stimuli.

References

Garner, W. R. Context effects and the validity of loudness scales. *J. exp. Psychol.*, 1954, **48**, 218–224.

Gebhard, J. W., Mowbray, G. H., & Byham, C. L. Difference-limens for photic intermittence. *Quart. J. exp. Psychol.*, 1955, **7**, 49–55.

Mowbray, G. H., Gebhard, J. W., & Byham, C. L. Sensitivity to changes in the interruption rate of white noise. *J. Acoust. Soc. Amer.*, 1956, **28**, 106–110.

Pollack, I. Auditory flutter. *Amer. J. Psychol.*, 1952, **65**, 544–554.

Reese, T. W. The application of the theory of physical measurement to the measurement of psychological magnitudes, with three experimental examples. *Psychol. Monogr.*, 1943, **55**(3, Whole No. 251).

Stevens, J. C. Stimulus spacing and the judgment of loudness. *J. exp. Psychol.*, 1958, **56**, 246–250.

Stevens, S. S. Concerning the form of the loudness function. *J. Acoust. Soc. Amer.*, 1957, **29**, 603–606. (a)

Stevens, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153–181. (b)

Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.*, 1957, **54**, 377–411.

Stevens, S. S., & Poulton, E. C. The estimation of loudness by unpracticed observers. *J. exp. Psychol.*, 1956, **51**, 71–78.

Stevens, S. S., Rogers, M. S., & Herrnstein, R. J. The apparent reduction of loudness: A repeat experiment. *J. Acoust. Soc. Amer.*, 1955, **27**, 326–328.

(Received December 8, 1958)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

TEMPORAL NUMEROSITY: IV. A COMPARISON OF THE MAJOR SENSES

CARROLL T. WHITE AND PAUL G. CHEATHAM

U. S. Navy Electronics Laboratory *Office of Naval Research*

In the first three papers of this series, studies were reported which dealt with the way in which rapidly presented sequences of visual and auditory stimuli are perceived. The first experiments were carried out with vision, and in these it was found that the number of flashes perceived followed a highly predictable pattern, a pattern determined to a greater extent by the time taken to present a sequence than by the number of stimuli presented in that sequence (Cheatham & White, 1952). Next, a study was made of the electrical activity of the retina during repetitive stimulation of the eye in order to determine whether or not some aspect of this activity might account for the limitation of perceptual rate which was found in the earlier studies (White, Cheatham, & Armington, 1953). The results of this work indicated rather clearly that the eye itself was not primarily responsible for this limitation, although it appeared very likely that some activity of the scotopic visual system was tending to decrease the perceived number of flashes.

The third study dealt with the perception of sequences of auditory stimuli (Cheatham & White, 1954). In this study it was quite obvious that perceived number was primarily a function of time, and that the upper limit of the perceptual rate was about 10 perceptual units/sec. Some very important questions still remained unanswered, however. In order to see if this phenomenon of perceptual limiting was to be found in senses other than vision and audition, the analogous study has been done with the sense of touch. The present paper reports the findings of this study, and in addition describes two visual experiments in which scotopic activity has been minimized.

METHOD

Tactual numerosity (Cond. 1).—In order to produce the stimuli for this study the same general setup was used that has been described in earlier reports in this series (Cheatham & White: 1952, 1954). A beam of light was brought to focus in the plane of a rotating disk, which was slotted to produce approximately square pulses of 10-msec. duration. A separate disk was used for each stimulus rate, 30/sec. and 10/sec., since the disks were driven by a constant-speed motor geared down to 1 rps. In this way it was possible to maintain a constant flash duration regardless of stimulus rate. A photoelectric tube was placed in the light path a few inches beyond the disk. Whenever light reached this tube the photocurrents were suitably amplified and led to an earphone. In order to make the resultant reaction in the earphone available to the tactile organs a thumb tack was cemented to the diaphragm and the point of the tack was then inserted under the side of *O*'s thumbnail.

No attempt was made to quantify the resultant stimulation. The gain was adjusted so that the vibrations were clearly perceptible to each *O*. The *O*'s hand which was being stimulated was wrapped in foam rubber so that the click of the earphone would not be heard. As a further precaution *O* was made to wear a pair of padded earphones. The *O*s were instructed to report the number of pulses perceived. The stimuli were presented by means of a modified method of limits, with five ascending and five descending number series for each stimulus rate. A ready signal was given 2 sec. before each series of pulses was presented. Five *O*s were used for this study, each making 10 judgments for each number-rate combination.

Visual numerosity, high surround (Cond. 2).—Light flashes were produced at the rate of 30/sec. by means of the slotted disk technique described above. After passing through the slots the light struck a piece of ground glass and was masked down to a spot ¼ in. in diameter. The brightness of the uninterrupted light stimulus was 600 ft.-L., and, as before, the duration of each flash was approximately 10 msec. On *O*'s side of the apparatus the stimulus spot was observed binocularly through a hole cut in white construction board. The brightness of this surround was maintained at 1000 ft.-L. by means of a small spotlight. In order to prevent the surround light from masking the stimulus, the ground glass was placed about 1 in. behind the plane of the surround. The *O* was seated so that his eyes were about 28 in. from the stimulus spot.

The *O*'s task was the same as in the previous studies. Upon presentation of a series of flashes he reported the number of flashes perceived. The stimuli were presented in alternating ascending and descending number series. The brightness of the surround made it impractical to have *O* maintain fixation throughout the experimental session, so he was instructed to avert his eyes until the ready signal was given. The 2-sec. interval between the ready signal and the stimulus was found to be ample time for *O* to fixate the target area and prepare for the stimulus. Four *O*s were used in this study, each making 15 judgments for each value of N_o.

Visual numerosity, strobe source, high surround (Cond. 3).—The special apparatus for this study consisted of a General Radio Strobotac and a Keystone Flashmeter. This combination proved to be quite flexible and well suited for work of this kind since various exposure times could be obtained with the Flashmeter and any desired flash rate could be obtained with the Strobotac. Practical considerations led to the choice of two exposure times (1/10 sec. and 1/5 sec.) and four flash rates (15/sec., 20/sec., 25/sec., and 30/sec.). The duration of each flash was about 5 μsec., as opposed to the flash durations of 10 msec. used in all the other visual experiments in this series.

The viewing arrangements for *O* were the same as those described in Cond. 2, including the 1000 ft.-L. surround. The Flashmeter was placed between the Strobotac and the ground glass so that only the light passed by the Flashmeter could reach the stimulus field. Each *O* was presented a separate series for both the exposure times, but within these two series the flash rates were varied at random until *O* had made 10 judgments for each rate. A ready signal was given 2 sec. before each stimulus presentation. Five *O*s were used in this study.

Results

The data obtained in Cond. 1 and 2 are shown in Table 1, and those for Cond. 3 are shown in Table 2. In Table 1 the distribution of perceived number is shown for each number of stimuli actually presented at the indicated presentation rates, 30/sec. and 10/sec. for the tactual study and 30/sec. only for the visual study.

TABLE 1

Distribution of Responses to Varying Numbers of Vibratory and Visual Stimuli at Various Rates of Presentation (Cond. 1 & 2)

Number of Stimuli (N_o)	Vibratory Stimuli															Visual Stimuli							
	30/sec. Rate Perceived Number (N_a)								10/sec. Rate Perceived Number (N_a)							30/sec. Rate Perceived Number (N_a)							
	1	2	3	4	5	6	7	*M*	1	2	3	4	5	6	*M*	1	2	3	4	5	6	7	*M*
1	50							1.0	50						1.0	60							1.0
2	3	47						1.9		50					2.0	57	3						1.1
3		29	21					2.4		6	44				2.9	10	42	8					2.0
4		8	39	3				2.9			19	31			3.6	4	31	25					2.4
5		1	35	14				3.3				34	16		4.3		8	42	10				3.0
6			21	29				3.6				4	42	4	5.0			45	14	1			3.3
7			7	32	11			4.1										21	31	8			3.8
8				25	22	3		4.6										4	42	14			4.2
9			1	11	32	6		4.9										3	26	26	5		4.6
10				2	28	19	1	5.4											12	34	13	1	5.1

TABLE 2

DISTRIBUTION OF RESPONSES TO 100-MSEC. AND 200-MSEC. EXPOSURES OF A STROBE LIGHT FIRING AT VARIOUS RATES, WITH OS ADAPTED TO A 1000 FT.-L. BRIGHTNESS LEVEL

Strobe Firing Rate	Mean Flashes/ Exposure (N_e)	Perceived Number (N_p)					
		1	2	3	4	5	Mean
100-Msec. Series							
15/sec.	1.5	21	28	1			1.6
20/sec.	2.0	10	31	9			2.0
25/sec.	2.5	5	24	21			2.3
30/sec.	3.0	1	27	18	4		2.5
200-Msec. Series							
15/sec.	3.0		24	24	2		2.6
20/sec.	4.0		3	41	7		3.1
25/sec.	5.0		1	24	24	1	3.5
30/sec.	6.0		1	19	19	11	3.8

Since the exact number of stimuli presented in the strobe-light study could not be controlled, Table 2 shows, instead, the average number which would be expected under each flash rate-presentation time combination, along with the distribution of judgments made for each such condition.

DISCUSSION

The most pertinent findings made to date in the series of studies on the topic of temporal numerosity are presented in Fig. 1. It is to be noted that all the data represented are for 30/sec. stimulus rates. The applicable data from the present work are included, along with portions of the data from previously published visual and auditory studies (Cheatham & White: 1952, 1954). The results of the three studies described in this paper go far in confirming the findings of the earlier work. The similarity of the auditory and tactual results is quite remarkable, for example, considering the differences between the conditions of the two experiments.

The marked difference between the curves representing the visual study performed with the eye adapted to a low brightness level and the visual study performed with the eye adapted to a high brightness level tends to verify the hypothesis that some characteristic of the rods was responsible for the type of function obtained in the earlier work (Cheatham & White, 1952). There is an indication, however, that even at the high adaptation level the eye responds more to the first flash than it does to subsequent flashes in a sequence, since there is a fusion effect to be seen at the start of the visual (high-surround) curve. Since this is probably a peripheral effect, we might be justified in ignoring the first flash when making a comparison of the visual results with the other two modalities. If this is done, it may be seen that the visual curve would match the auditory and tactual curves almost exactly. Since we are more interested in perceptual rate than in absolute values, however, it isn't really necessary to do this, since it can be clearly seen that the slopes of all three curves are practically identical.

In the strobe-light experiment (Cond. 3) it was possible to obtain only two

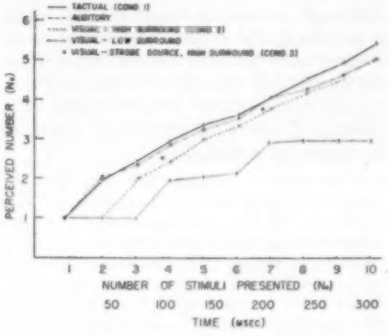

FIG. 1. Perceived number as a function of the number of stimuli presented, for the visual, auditory and tactual modalities. The data for the visual–low surround and the auditory curves have been taken from two earlier papers in this series (Cheatham & White: 1952, 1954). All curves represent stimulus presentation rates of 30/sec.

points which could be entered with the group of 30/sec. curves. These two points, however, fit very well with the other results, showing that stimulus duration is not a very important factor in numerosity judgments. The slope of the line determined by these points is seen to be about the same as that for the auditory, tactual, and light-adapted visual curves—approximately 80 msec./perceived unit (12/sec.). It should be noted that this value is derived from relatively short sequences of stimuli. In the work on auditory numerosity (Cheatham & White, 1954), when longer sequences were involved, the corresponding value was estimated as being nearer to 90 msec./perceived unit (11/sec.).

The results of this series of studies clearly indicate that we are dealing with some very basic aspect of the functioning of the nervous system, but the nature of the experiments has not been such that we can infer just what is involved. These findings bear directly on certain ideas concerning CNS functions which have been stated by Wiener (1948, p. 165); and some of these data have been used to illustrate the even more explicit "moment function hypothesis" advanced by Stroud (1956). Although we do not know exactly what is involved, the results of this series of studies seem to lend strong support to the hypothesis that *there is some temporal process in the central nervous system that limits and orders the perceptual events of the major sense modalities.*

Summary

Studies were carried out to determine the perceived number of vibratory stimuli as a function of the number presented for rates of 10/sec. and 30/sec., and data were obtained on the visual perception of number when the eye was adapted to a high light level. The results of these studies were compared with the results of previous work on vision and audition and it was found that the upper perceptual rates for all three senses were approximately the same, about 80 msec./perceived unit. This is interpreted as further evidence in favor of the hypothesis that there is some temporal process in the central nervous system that tends to limit the perception of the inputs of the major sense modalities.

References

CHEATHAM, P. G., & WHITE, C. T. Temporal numerosity: I. Perceived number as a function of flash number and rate. *J. exp. Psychol.*, 1952, **44**, 447–451.

CHEATHAM, P. G., & WHITE, C. T. Temporal numerosity: III. Auditory perception of number. *J. exp. Psychol.*, 1954, **47**, 425–428.

STROUD, J. M. The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology.* Glencoe: Free Press, 1956. Pp. 174–207.

WHITE, C. T., CHEATHAM, P. G., & ARMINGTON, J. C. Temporal numerosity: II. Evidence for central factors influencing perceived number. *J. exp. Psychol.*, 1953, **46**, 283–287.

WIENER, N. *Cybernetics.* New York: Wiley, 1948.

(Received December 9, 1958)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

THE VON RESTORFF ISOLATION EFFECT: TEST OF THE INTRALIST ASSOCIATION ASSUMPTION [1]

ELI SALTZ AND SLATER E. NEWMAN

Wayne State University *North Carolina State College*

The von Restorff isolation phenomenon refers to the rapid learning of a dissimilar item in the midst of a serial or paired-associates list of similar items—e.g., the rapid learning of a nonsense term in the midst of a serial list of nouns. The best verified aspect of the von Restorff isolation phenomenon in *serial* learning is the rapidity with which the isolated term is learned as a *response*. In multitrial situations, such results have been reported by Kimble and Dufort (1955), Newman and Saltz (1958), Smith and Stearns (1949), and von Restorff (1933) herself, among others. It is important to note, however, that strong tendencies to recall the isolated term are also found after a *single* presentation of the serial list. Such results have been reported by Green (1956), Postman and Phillips (1954), Smith (1949), and von Restorff (1933). Thus, apparently, the mechanisms basic to the von Restorff effect begin to operate after only one presentation of a serial list. For this reason, the single trial situation appears to be very important for investigating the basic mechanism of the von Restorff effect. The present experiment is primarily concerned with the isolation phenomenon as it occurs after a single learning trial. The purpose of this experiment is to determine more about the characteristics of the isolation phenomenon, and to examine the adequacy of the prevalent class of S-R theories, used to explain this phenomenon.

To the writers' knowledge, no S-R psychologist has proposed that different mechanisms are basic to single and multitrial isolation phenomena. In general, the assumption behind most S-R theories of the von Restorff effect appears to be that isolation results in a relatively stronger *bond* between the isolated term and its appropriate stimulus or response term within the list. For convenience, we shall refer to this class of explanations a *intralist S-R association theories*. S-R theorists have tended to differ mainly in the type of mechanisms postulated to produce such a strengthening of the S-R bond. Gibson (1940), for example, assumes that, due to intralist stimulus generalization, responses will at times be evoked by stimuli other than those to which they are correct. The difficulty of learning a given S-R pair will be a function of the number of such incorrect responses attached to the stimulus item. Since the isolated term is dissimilar from the other items in the list, there should be less generalization of incorrect, competing responses to it than to the other items in the list, and thus the effective strength of the bond between the isolated item, as a stimulus, and its correct response will be relatively great. It should be noted that since Gibson's theory is concerned with stimulus generalization it cannot, in its present form, be used to account for the rapid learning of the isolated term as a *response*. Gibson's theory can be used to account for the rapid learning of an S-R pair in which the isolated term is the stimulus; however, this effect has not been clearly demonstrated in serial learning studies. Newman and Saltz (1958) have used the response generalization concept in an attempt to handle the rapid learning of the isolated item as a response. Response generalization

[1] The present research was supported in part by a grant from the Wayne State University Graduate School. Preliminary portions of the research were conducted under ARDC Project No. 7709, Task 37304, in support of the research and development program of the Maintenance Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Colorado. The opinions expressed are those of the authors, and do not necessarily represent the views of the United States Air Force.

is the phenomenon that occurs when a stimulus which is associated with Response A evokes Response B because of the similarity that exists between the two responses. Since the isolated term is dissimilar from the other list items, the tendency of the stimulus for the isolated term to evoke incorrect responses due to response generalization should be reduced. This again is an intralist S-R association theory, since the predication of an association between the isolated item and its intralist stimulus term is basic to it. Kimble and Dufort (1955) also develop a theory about the manner in which the isolation of a term might lead to a facilitation of the connection between the isolated term and its appropriate stimulus term.

Since most multitrial studies have used an anticipation learning technique, rapid learning of the isolated item has meant correct elicitation of the isolated item by its appropriate stimulus term. Thus, at an *empirical* level, it seems safe to say that isolation facilitates intralist associations in multitrial learning. However, there is some evidence to indicate that, even at the empirical level, such facilitation of intralist associations does not occur after a *single* presentation of the isolated list. In the Newman and Saltz (1958) multitrial isolation study, it was possible to examine the number of correct anticipations of the isolated term as a response during the trial following the first presentation of the serial list. These results did not indicate a significant advantage for the correct elicitation of the isolated responses over their nonisolated counterparts in the control lists (though the isolated responses occurred as *intrusion errors* significantly more frequently than the control items).

While the Newman and Saltz results for number of correct responses on the first test trial do not support expectations from the intralist association assumption, these results are far from conclusive since very few responses, correct or incorrect, were given during the crucial first trial. The *Ss* frequently suppress responses when they are uncertain about their correctness. Consequently, the present study represented an attempt to obtain a more sensitive measure of associative strength.

In the present experiment, the isolation effect was measured by means of a controlled associations test. After a single learning trial, *Ss* were shown the items of the list in a random order and were asked to give the first word from the list that came to mind upon seeing each word. If isolation facilitates formation of associations, then in a controlled association task the isolated term should occur as a response to its appropriate stimulus term more frequently than does its nonisolated counterpart in a control list. Control groups tested by means of free recall were also run to insure that the material used would produce the typically reported results.

The hypotheses to be tested in the present paper arise from the S-R association class of explanations for the von Restorff effect.

1. According to this class of explanations, isolation of an item in a list increases the effective strength of the connection between the isolated term and its intralist stimulus. Therefore, in the controlled association test the isolated item should be elicited by its appropriate intralist stimulus term more frequently than the control item is elicited by its appropriate stimulus.

2. In terms of response generalization theory, the tendency for one response to intrude on another is assumed to be a function of the similarity between the responses. Since the nonisolated control item is similar to the other items of the list, while the isolated item is not, the isolated item should occur *less* frequently than the control item as an intrusion to inappropriate intralist stimuli.

The serial lists employed in the present study were identical to those previously used by Kimble and Dufort (1955) and by Newman and Saltz (1958): a nonsense word embedded in a serial list of nouns. Thus, meaningfulness and isolation were confounded in these studies. Such confounding might be expected to reduce the likelihood of the isolated term being learned rapidly, since nonsense terms

are more difficult to learn than meaningful words. From this point of view, the use of a nonsense term allowed for a more conservative test of the isolation effect. If an intralist S-R association theory is to be seriously offered as an explanation for the studies in question, a nonsense term as the isolated item would be expected to fulfill the relationships specified in the hypotheses outlined above.

Method

Subjects.—The experiment was originally conducted using, as *Ss*, 80 airmen awaiting training in Air Force technical schools. The experiment was replicated at Wayne State University using 80 introductory psychology students as *Ss*.

Materials.—Two of the serial lists previously used by Kimble and Dufort (1955), and also by Newman and Saltz (1958), were used in the present experiment. These two lists consisted of 13 items each. Of these 13 items, the two lists had 12 nouns in common; these 12 were from the high meaningfulness extreme of Noble's (1952) meaningfulness scale. The item in the seventh serial position was different for the two lists; in the "nonisolated" control list this item was "office," also from the high end of Noble's meaningfulness scale, and in the "isolated" list this item was "gojey," a paralog of low meaningfulness in Noble's scale.

Procedure.—The two dimensions of the experimental design were: (*a*) list to be learned (isolation list vs. nonisolation list), and (*b*) method of testing for learning (free recall vs. controlled associations). The 80 *Ss* in each replication were randomly assigned to the four groups, 20 *Ss* per group.

Each *S* was tested individually. Prior to the single learning trial, each *S* was instructed to learn the list by means of the serial anticipation method. The *Ss* were not told that they would receive only one learning trial. In the controlled association tests, *Ss* were next shown the list items in one of four random orders and were asked to give the first word from the list that came to mind upon seeing each item. In the recall test, *Ss* were asked to recall as many of the list items as they could.

All items were presented for learning on a memory drum. Original learning was at a 2-sec. rate. The controlled associations test was presented on 3×5 cards and *Ss* were permitted sufficient time for a response.

Results

The first issue to be dealt with was the occurrence of the isolation effect in the recall conditions. As anticipated from previous studies, the isolated term was *recalled* significantly more frequently than its nonisolated counterpart. In the Air Force replication, 16 *Ss* (80%) in the Isolated-Recall group emitted the isolated term; only five (25%) in the Nonisolated-Recall group emitted the control item. In the Wayne replication the results were almost identical: 17 *Ss* (85%) in the Isolated-Recall group emitted the isolated term; 6 *Ss* (30%) in the control group recalled the control item. In each replication these results are significant at well beyond the .01 level. These results indicate that the materials used were susceptible to the von Restorff recall effect.

Next the results of the controlled association tests were evaluated. In the Air Force replication, only three (15%) of the *Ss* gave the isolated item as a correct response to its appropriate stimulus; four *Ss* (20%) in the control group gave the control item as a correct response to its appropriate stimulus term. Similarly, in the Wayne replication, three *Ss* (15%) gave the isolated item as a correct response while two *Ss* in the control group gave the control item. These results do not support the hypothesis that isolation facilitates the association between the isolated item and its appropriate stimulus, after a single presentation of the serial list.[2]

[2] The isolation effect in recall was significantly greater than in controlled associations beyond the .05 level when evaluated by means of the test for heteregeneity described by Snedecor (1956). This test involves

Thus, for the case involving a single learning trial, hypotheses such as those suggested by Gibson (1940) and Kimble and Dufort (1955) appear to be inapplicable.

In a previous study (Newman & Saltz, 1958), the writers found that the isolated item occurred as an intrusion error in the first anticipation trial of serial learning significantly more frequently than did the control item. It would be anticipated, therefore, that similar results would occur in the controlled associations data of the present experiment. This expectation is borne out. The isolated item was elicited as a response to stimuli other than the appropriate stimulus by 13 *Ss* (65%) in the Air Force group, 11 *Ss* (55%) in the Wayne group. In the control groups, on the other hand, the control item was elicited as a response to stimuli other than the appropriate stimulus by 6 *Ss* (30%) in the Air Force group, 7 *Ss* (35%) in the Wayne group. Over the pooled Air Force and Wayne groups, the frequency of inappropriate elicitation of the isolated term was significantly greater than the frequency of inappropriate elicitation of the control term at close to the .01 level. The interaction between isolation effects and replication, evaluated by means of Snedecor's (1956) test for heterogeneity, was far from significant.

Since the isolation effect appeared strongly in the recall data, these results were analysed in greater detail. Superficially, at least, responses emitted during the recall test appear to be less bound to intralist stimuli than the responses emitted during controlled association test. However, the item recalled immediately before the recall of the isolated term may act, in part at least, as a stimulus to the isolated term. Consequently, the recall data were analysed to determine possible stimulus effects. The Air Force and Wayne groups were combined for purposes of this analysis, since the number of *Ss* in the control groups who emitted the control item was too small to permit separate analysis for each replication. This appears to be legitimate since the replications produced very similar results. For this analysis, only those *Ss* were considered who, in the test trial, recalled *gojey* (from Isolation-Recall group) or *office* (from Nonisolation-Recall group). For each of these *Ss*, the item emitted immediately prior to the recall of *gojey* or *office* was considered the stimulus term. The serial position distance in the original learning list was then computed, for each *S*, between the response term (*gojey* or *office*) and the stimulus term (the term recalled immediately prior to *gojey* or *office*). Direction of association was ignored; e.g., a backward association of one serial position was considered equivalent to a forward association of one serial position.

Examination of these data indicates that the control item, *office*, is extremely stimulus bound. For 40% of the *Ss*, the control item was preceded in recall by the item that immediately preceded it or immediately followed it in the original list (i.e., a serial position distance of one position); for 60% of the *Ss*, the control item was preceded by a term within two serial positions distance in the original list. As can be seen from these data, the distribution of serial distances is markedly positively skewed for the control group.

The results for the isolated group are quite different, being markedly negatively skewed. The isolated item was preceded in recall by the item

evaluation of the interaction between two 2×2 tables.

that immediately preceded it or immediately followed it in the original serial list by only 14% of the *Ss*. Only for 24% of the *Ss* was the isolated item preceded in recall by an item within two serial positions in the original list. The difference in percentages of *Ss* for whom the recall stimulus term was within two serial positions distance in the original list is significant at beyond the .05 level of confidence when evaluated by means of McNemar's (1949) Formula 27a.[3]

The serial position distance data also fail to support the intralist association assumption, since the control item appears to be determined more strongly than is the isolated item by the prior emission of a term that was close to it (in terms of serial position distance) in original learning. It might be argued that if the association between the isolated term and its appropriate stimulus were very strong, this might generalize to other stimuli, so that many items in the list might be capable of eliciting the isolated term. Such an eventuality could conceivably result in the serial distance data reported above. However, this argument is not supported by the association data reported in the main body of this experiment. A generalization explanation is dependent on a strong connection between the isolated term and its appropriate stimulus; the association data fail to indicate such a strong association.

On the basis of the results so far reported, one might expect a tendency for the isolated term to be the first item recalled. This was true for only two *Ss* in each of the replications, or 10% of all *Ss* in the Isolation-Recall groups. In the two replications combined, 20 *Ss* (50%) emitted the first word in the original learning list as their first recall item. Both the isolated and control items tended to be the fifth word emitted during recall: For the Wayne group, the mean serial position of recall for both the isolated and control group was identical, 4.8. For Air Force groups the means were 4.5 and 4.8, respectively. Thus, despite the fact that the isolated word was recalled more frequently than the control word, isolation did not result in early emission of the isolated term. This result appears to be contrary to expectations from von Restorff's (1933) "figure-ground" theory and Green's (1956) "surprise" hypothesis. However, these theories have not been enunciated in sufficient detail to permit unequivocal predictions on this point.

On the basis of Gibson's (1940) theory one would predict that isolation will facilitate the connection between the isolated item and its appropriate intralist response. Examination of the data fails to support this prediction: In neither recall nor controlled association was there even a tendency in favor of the Gibson prediction. Thus, isolation facilitated learning the isolated item as a response, but did not facilitate learning the appropriate response to the isolated term as a stimulus. These results are consistent with the general body of results in serial learning studies: e.g., Kimble and Dufort

[3] The writers' original intention was to compare percentages of *Ss* for whom the *gojey* or *office* response was *one* serial position distance from the recall-test stimulus—i.e., for whom stimulus and response were adjacent. However, this resulted in cell frequencies which were too small for the analysis. It should be pointed out that five *Ss* gave *office* or *gojey* as their first recall response; thus these *Ss* had no recall-test stimuli for *office* or *gojey* and were omitted from the analysis.

A similar analysis was conducted for the controlled association data. While the results were in the same direction as the recall results, they did not reach an acceptable level of significance.

(1955) and Newman and Saltz (1958), using the same materials as used in the present study; Jenkins and Postman (1948), in which a meaningful word was embedded in a serial list of nonsense syllables; Jones and Jones (1942), in which isolation was produced by embedding a syllable typed in red within a list of syllables typed in black. Thus these results do not appear to be a peculiarity of the particular materials used in the various studies.

Discussion

On the basis of the present study, the writers' (Newman & Saltz, 1958) previous study, and the body of literature in this area, the following empirical conclusions seem justified concerning the effects of isolation in serial learning. (*a*) The primary effect of isolation appears to be an accelerated learning of the isolated item as a response. (*b*) After a single presentation of the serial list, the isolation effect manifests itself in terms of increased recall of the isolated term in a recall test, and in terms of frequent elicitation of the isolated term by inappropriate stimuli in controlled association or anticipation tests. The effect does not manifest itself as an increase in correct elicitation of the isolated term by its appropriate stimulus in either controlled association or anticipation tests as contrasted with its nonisolated control list. In fact, there is evidence to suggest that isolation actually decreases the dependence of the isolated term on its appropriate stimulus in recall. (*c*) After several learning trials a tendency develops for the isolated term to be correctly anticipated by its appropriate stimulus.

The discrepancy in results between isolation effects in single-trial situations and multitrial situations raises a crucial question: Are the same fundamental mechanisms involved in both types of studies? There is no way of determining this at present, of course. It would be unexpected, however, to find that two unrelated mechanisms are required to explain the data when the identical manipulation leads to the same type of behavior (i.e., increased learning of the isolated item as a response) in two situations that differ only in number of learning trials and type of response measure. As has been pointed out previously, the von Restorff effect does occur after a single learning trial when emissions (i.e., recall) are used as a criterion; thus it seems reasonable to consider the multitrial effect as somehow developed on the basis of mechanisms already operative during the first trial. If this proves to be the case, it would be anticipated that some assumption other than that of intralist association would be required for explaining the von Restorff effect in multitrial learning. Indeed, in a previous paper, Newman and Saltz (1958) found serious inadequacies in the intralist S-R strength theory most widely held to date—the stimulus and response generalization theory of isolation.

The results of the present study are not consistent with the usual conceptions of response generalization. Since the nonisolated control item is similar to the other items of the list, one would expect it to occur more frequently than the isolated term as an intrusion error. The results, however, are opposite to such an expectation. Since generalization in verbal learning is typically defined in terms of intrusion errors, the writers believe that these results indicate that there is still much to be learned about the functioning and interaction of response generalization and similarity as these influence intrusion tendencies.

Finally, it should be emphasized that the present study was not designed to support a new theoretical position concerning the von Restorff effect. It was designed to investigate characteristics of the isolation effect in serial learning. The results suggest that the effect is not what it has been assumed to be. The effect appears, from the present study, to be an increased tendency for the

isolated term to be *emitted*, without being strongly tied to its intralist stimulus term. These results indicate that previous S-R theories are not adequate to handle the phenomenon.

Summary

The best verified aspect of the von Restorff isolation phenomenon in serial learning is the rapidity with which the isolated term is learned as a response. This effect has been reported in both multitrial learning and after a single presentation of the serial list. Since it appears reasonable to assume that the multitrial effect is based on the variables determining the single trial effect, the present study investigated the single trial effect.

The present paper was designed to test the assumption, implicitly made by most S-R theorists, that isolation facilitates the intralist association of the isolated term with its appropriate stimulus term.

Comparison of controlled association vs. recall data after a single presentation of the list, indicated: isolated terms do not occur as correct responses significantly more frequently than nonisolated terms, but are emitted significantly more frequently than nonisolated terms. Results were interpreted as indicating that intralist facilitation theories, as a class, are probably inappropriate for explaining the von Restorff effect.

References

Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.

Green, R. T. Surprise as a factor in the von Restorff effect. *J. exp. Psychol.*, 1956, **52**, 340–344.

Jenkins, W. O., & Postman, L. Isolation and "spread of effect" in serial learning. *Amer. J. Psychol.*, 1948, **61**, 214–221.

Jones, F. M., & Jones, M. H. Vividness as a factor in learning lists of nonsense syllables. *Amer. J. Psychol.*, 1942, **55**, 96–101.

Kimble, G. A., & Dufort, R. H. Meaningfulness and isolation as factors in verbal learning. *J. exp. Psychol.*, 1955, **50**, 361–368.

McNemar, Q. *Psychological statistics*. New York: Wiley, 1949.

Newman, S. E., & Saltz, E. Isolation effects: Stimulus and response generalization as explanatory concepts. *J. exp. Psychol.*, 1958, **55**, 467–472.

Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.

Postman, L., & Phillips, L. W. Studies in incidental learning: I. The effects of crowding and isolation. *J. exp. Psychol.*, 1954, **48**, 48–56.

Smith, M. H., Jr. The influence of isolation on immediate memory. *Amer. J. Psychol.*, 1949, **62**, 405–411.

Smith, M. W., Jr., & Stearns, E. G. The influence of isolation on the learning of surrounding materials. *Amer. J. Psychol.*, 1949, **62**, 369–381.

Snedecor, G. W. *Statistical methods* (5th ed.) Ames: Iowa State Coll. Press, 1956.

von Restorff, H. Ueber die wirkung von bereichsbildungen im spurenfeld. In W. Kohler and H. von Restorff, *Analyse von vorgangen im spurenfeld*. I. *Psychol. Forsch.*, 1933, **18**, 299–342.

(Received October 6, 1958)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

INTENTIONAL AND INCIDENTAL LEARNING UNDER HIGH AND LOW EMOTIONAL DRIVE LEVELS

DONALD H. KAUSLER, E. PHILIP TRAPP, AND CHARLES L. BREWER[1]

University of Arkansas

Bahrick (1954) has recently demonstrated that the amount of incidental learning is inversely related to the strength of the incentive associated with the intentional learning task. Manipulating a well-defined incentive condition, presence or absence of a monetary reward, he found that high incentive *Ss* displayed significantly greater intentional learning but significantly less incidental learning than low incentive *Ss*. These results were in accord with his hypothesis that increasing the goal-demand should decrease incidental learning because the set relating to irrelevant components of the task would have less success competing against the stronger set directed toward the relevant components. A similar interpretation has been proposed by Bruner, Matter, and Papanek (1955).

The present study is concerned with the effects upon intentional and incidental learning of motivation as manipulated by an emotionally based drive, anxiety. In the Bahrick situation motivation is clearly manipulated by an incentive oriented set. Anxiety, however, represents a generalized drive which conceivably may have energizing (e.g., Spence, 1958) and/or interference effects through the elicitation of task-irrelevant responses (e.g., Child, 1954). In order to compare directly the effects of incentive defined drive and emotionally based drive on intentional and incidental learning, the learning tasks used in this study duplicate the learning tasks employed by Bahrick. This procedure recognizes the importance of task characteristics as variables in learning studies. In an earlier study, Silverman and Blitz (1956) found that high drive *Ss*, as defined by score on the Taylor MAS (Taylor, 1953), did not differ from low drive *Ss* in intentional learning but did display significantly less incidental learning. However, a meaningful comparison between anxiety- and incentive-defined motivation in terms of intentional-incidental learning effects is made difficult by the grossly different tasks, both intentional and incidental, employed in the two studies (e.g., serial geometric forms in the Bahrick study and serial nonsense syllables in the Silverman and Blitz study as intentional tasks).

Two separate experiments are reported here. In Exp. I, the emotional drive is intrinsic-drive level as measured by the Taylor MAS. In Exp. II, the emotional drive is extrinsic or induced-anxiety induced by ego-involving instructions. Thus Exp. II serves to extend further the comparison between incentive-oriented and emotionally based motivation by distinguishing between *E*-induced anxiety and *S*-measured manifest anxiety.

EXPERIMENT I

Method

Subjects.—The *Ss* consisted of 50 students[2] in courses in abnormal psychology at the University of Arkansas, representing approximately the top quartile (high anxiety, MAS

[1] Now at the College of Wooster.

[2] One *S* in the low anxiety quartile had to be disqualified for deliberately learning the incidental material.

score of 19 and above, $N = 26$) and the bottom quartile (low anxiety, MAS score of 9 and below, $N = 24$) for the distribution of 98 students on the Taylor MAS. The high anxiety group (Group HA) contained 13 men and 13 women; the low anxiety group (Group LA) contained 14 men and 9 women.

Learning tasks.—Both the nature and presentation of the learning materials duplicated the Bahrick (1954) study. The intentional task involved learning a serial list of 14 geometric forms (seven different forms, such as square, triangle, moon, etc., each appearing twice in the sequence). The learned responses consisted of the anticipation of the next form in the series. Each form was drawn so as to occupy an area of approximately ¼ sq. in. and filled with one of seven colors. Thus each of the seven different forms had associated with it two different colors. Incidental learning was measured by showing *Ss* a 14-row by 7-column chart with the forms along the rows in the same sequence as in the learning trials and the colors along the columns. The *Ss* checked the color associated with each of the 14 forms in the series.

Procedure.—The Taylor MAS, labeled "Biographical Inventory," was administered during regular class periods. The *Ss* were told that the information was collected to compare students at the University of Arkansas with students elsewhere. The MAS, in an abbreviated form, contained the 50 anxiety keyed items and 40 buffer items. Two to four weeks later *Ss* in the two extreme quartiles on the MAS were asked individually by their instructor to participate in a learning experiment. The *Ss* then learned the material individually with instructions to learn the series of geometric forms. The colors were not mentioned in these instructions. Bahrick's motivational instructions, of course, were not employed. Apparently none of the *Ss* suspected any relationship between the two parts of the experiment.

Results and Discussion

Two measures of intentional learning were obtained for each *S*: trials to a criterion of one perfect trial (T) and total number of errors made to reach the criterion (E). Incidental learning was measured by the number of correct color recognition (C). Table 1 shows the means and *SD*s of the two groups on T, E, C, and C′ (adjusted number of correct color recognitions where the adjustment was made on the basis of the regression of C on T).

TABLE 1

SUMMARY DATA FOR HIGH AND LOW INTRINSIC ANXIETY GROUPS ON INTENTIONAL AND INCIDENTAL LEARNING

Measure	High Anxiety		Low Anxiety	
	Mean	*SD*	Mean	*SD*
T	18.7	5.67	24.3	8.99
E	80.42	29.47	114.13	63.76
C	6.69	2.85	5.65	2.59
C′	6.62	—	5.73	—

To test for a difference in intentional learning, Groups HA and LA were compared on both T and E. The distributions of scores for both groups were of sufficient skewness (positive direction) to warrant the use of a nonparametric test, the Mann-Whitney *U* test. For T, the difference between the two groups was significant at the .02 level ($z = 2.39$, $P < .02$, two-tailed test). For E, the difference between the two groups was significant at the .04 level ($z = 2.11$, $P < .04$, two-tailed test). The high emotional drive *Ss* were superior to the low emotional drive *Ss* on both T and E. This finding is comparable to Bahrick's finding of superiority for his high incentive group on the intentional task.

It is very unlikely that sex and intelligence could account for these differences between Groups HA and LA. A comparison of scores failed to reveal any significant difference between sexes for either learning measure. No significant difference was found in ACE scores between the two groups. Moreover, the correlation (rho) between ACE score and trials to the criterion was found to be − .04. In addition, the rho between ACE

score and color recognitions was only −.18.

To test for a difference in incidental learning, Groups HA and LA were compared on C. To control for a possible effect of T on C, an analysis of covariance was made. An insignificant F was obtained ($F = .89$, $df = 1/46$, Between-groups $MS = 8.45$, Within-groups $MS = 9.45$). Unlike the increased incentive condition in the Bahrick study, increasing emotional drive level did not result in a decrease in incidental learning. In fact, as seen in Table 1, there is a slight trend in the reverse direction. Thus the high and low drive groups appeared to be from the same population in regard to incidental learning scores.

Experiment II

Method

Subjects.—The *Ss* consisted of 50 students, 29 men and 21 women, in courses in general psychology at the University of Arkansas. All *Ss* represented volunteers for a learning experiment. They were assigned randomly to either a high motivation group (Group HM, $N = 25$) or a low motivation group (Group LM, $N = 25$).

Procedure.—The learning materials and method of presentation again duplicated the Bahrick study.

Following the reading of the general instructions, Group HM was given additional instructions intended to produce a high degree of ego involvement and induced anxiety. The instructions were:

"The study in which you are participating is a broad one involving a comparison between students in state universities in different parts of the country on learning tasks of varying difficulty. The task you are to learn is a complex one, involving certain aspects of general intelligence. Since University of Arkansas students in the past have usually performed poorer than students from state universities in other sections of the country on similar types of tasks, we are most hopeful that you will make an all-out effort to do your very best, so that Arkansas students will not again score below the national average when the results are published."

The additional instructions for Group LM were identical to those given by Bahrick to his low incentive group. These instructions were intended to induce a low degree of anxiety in *Ss*.

Results and Discussion

The same measures of intentional and incidental learning obtained in Exp. I were obtained. Table 2 shows the means and *SD*s of the two groups on T, E, C, and C′.

To test for a difference in intentional learning, Groups HM and LM were compared on both T and E. Since the scores for these groups were also skewed, the Mann-Whitney U test was again used to test for differences between groups. For T, the differences between groups was significant at the .06 level ($z = 1.88$, $P < .06$, two-tailed test). For E, the difference between groups was significant at the .02 level ($z = 2.40$, $P < .02$, two-tailed test). As in Exp. I, the high drive group excelled the low drive group, thus contributing additional evidence for the superiority of high emotional drive in the intentional learning of this task.

An analysis of covariance on C for the two groups again revealed no significant difference in incidental learning between the two emotional drive levels. The obtained F of .14 ($df = 1/47$, Between-groups $MS = .88$, Within-groups $MS = 6.42$) was obvi-

TABLE 2

Summary Data for High and Low Induced Anxiety Groups on Intentional and Incidental Learning

Measure	High Anxiety		Low Anxiety	
	Mean	*SD*	Mean	*SD*
T	20.3	8.32	24.2	7.84
E	85.00	42.91	111.93	44.30
C	6.44	2.45	6.56	2.50
C′	6.36	—	6.64	—

ously not significant. For induced emotional drive level, as well as for intrinsic drive level, high and low drive groups appeared to be from the same population in regard to incidental learning scores.

Discussion

In both experiments the high drive group was superior to the low drive group both in the number of errors committed and in trials to reach the criterion. These results were thus comparable to those found by Bahrick on the same task when motivation was manipulated by a positive incentive. In contrast to Bahrick's finding, however, our high drive groups did not display less incidental learning than the low drive groups. Apparently, new factors are introduced to confound and complicate the relationship when emotionally based drives are involved. It appears that the increased strength of the set toward the goal task does not take place to the detriment of the set toward irrelevant learning for *Ss* operating under higher levels of emotional drive. This suggests that emotionally based drives are more situationally oriented than nonemotional drives and less confined to the specified task. That is, the increase in drive level is expended on all components of the learning situation. Clearly, additional research is needed to test further this proposed differential between incentive oriented and emotionally based motivation.

On the other hand, there is another factor operating in the Bahrick study that might account for the difference between his results and those of the present study. The smaller amount of incidental learning in his monetary incentive group could simply be reflecting a decrease in motivation at the time of measuring incidental learning. The monetary incentive was specifically related to the intentional task of learning the list of forms. The absence of an incentive for performance in the incidental part of his study could conceivably engender a "let down" attitude with subsequent performance decrement. In our experiments, as a point of difference, there is no withdrawal of incentive in the high drive groups following intentional learning.

Summary

Two experiments were reported relating emotional drive level to intentional and incidental learning. In Exp. I, the drive manipulated was intrinsic, i.e., score on the Taylor MAS. In Exp. II, drive was induced by differential instructions. In both experiments the learning tasks replicated a study by Bahrick in which a high incentive group displayed significantly greater intentional learning but significantly less incidental learning than a low incentive group.

In both experiments the high drive group was superior to the low drive group in intentional learning but did not differ in incidental learning. The results suggested that emotional motivation is more situationally oriented than incentive oriented motivation and less confined to the relevant task.

References

Bahrick, H. P. Incidental learning under two incentive conditions. *J. exp. Psychol.*, 1954, **47**, 170–172.

Bruner, J. S., Matter, J., & Papanek, M. L. Breadth of learning as a function of drive level and mechanization. *Psychol. Rev.*, 1955, **62**, 1–10.

Child, I. L. Personality. *Ann. Rev. Psychol.*, 1954, **5**, 149–171.

Silver...n, R. E., & Blitz, B. Learning and two kinds of anxiety. *J. abnorm. soc. Psychol.*, 1956, **52**, 301–303.

Spence, K. W. A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

Taylor, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

(Received December 18, 1958)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

THE ROLE OF PAST EVENTS IN PROBLEM SOLVING[1]

JACQUELINE J. GOODNOW, IRVIN RUBINSTEIN, AND BETTY L. SHANKS

Walter Reed Army Institute of Research

How are we to measure the effect of some particular condition—age, anxiety, medication, or the like—on the process of problem solving? One way that has been suggested recently is to consider the organization or patterning of a series of steps over time, particularly the patterning that occurs in the relationships between a present step and the events that preceded it. Grant (1954) has discussed the possibilities of such an approach, and Frick (1953) has provided an excellent example demonstrating the disruptive effects of anxiety. In general, however, the approach has been used far less often than one would expect.

The present study applies this approach to two conditions—sleep loss and electroconvulsive shock (ECS). With both of these conditions, it has been suggested that changes occur in the role of past events but the changes have been difficult to document. With sleep loss or any condition like fatigue, Welford (1953) has proposed that *Ss* use fewer past events than is usual. For ECS, Janis (1950) has made the observation that *Ss* given ECS find it difficult to maintain a task-set. In both cases, the *Ss* seem to be less capable of tying a variety of past events into their actions or hypotheses.

To test this possibility, a useful task is one which requires that behavior vary to fit with accumulating information. With such a task, sleep loss or ECS should produce simpler or more fixed behavior based on the individual's ignoring part of the information provided by a large number of past events.

[1] We are deeply indebted to Ardie Lubin for both comment and advice.

General Method

Design.—In both Exp. I (sleep loss) and Exp. II (ECS), performance can be compared against two standards: the *S*'s own performance before and after the critical treatment, and the performance of control *Ss*. The specific occasions for testing are given in the descriptions of each experiment.

Task and instructions.—In both studies, *S* played a modified slot machine, a more automatic version of the machine described by Bush and Mosteller (1955), and by Goodnow (1955). On each of a series of trials, *S* chose between one of two keys. On each trial, the machine was programmed to pay off on the left-hand key or the right-hand key, never neither or both, and *Ss* were told this. The machine delivered a chip if *S* chose the correct key for that particular trial. In both studies, *Ss* were told that there was a pattern to the payoffs but that it was not possible, without the use of pencil and paper, to win a chip every time. The *Ss* were to aim at winning as many as possible. In the sleep loss study, where *Ss* played against a 50:50 random schedule, *Ss* were further advised that winning 50% of the time was "satisfactory," but barely so. In the ECS study, the machine paid off on a 67:33 schedule, with the left key scheduled to pay off 67% of the time. No information about a satisfactory level of success was given.

Analysis.—Hyman and Jenkin (1956) have proposed the use of multiple phi $\left(\sqrt{\frac{\chi^2}{n}}\right)$ to describe the degree of relationship between the events on one trial and the response on the next. For each *S*, chi square is computed from a contingency table (Table 1 is an example). The value of phi for Table 1 is .56. The maximum possible value of phi is 1, the minimum zero.

Essentially, phi reflects the extent to which the contingent responses shown in Table 1 describe something about *S*'s way of responding which cannot be told from the over-all

TABLE 1

CONTINGENCY TABLE SHOWING THE FREQUENCY WITH WHICH *S* CHOOSES ALTERNATIVES AFTER EACH OF FOUR PRECEDING EVENTS

Response on Trial *n*	Preceding Event Trial *n* − 1				Total
	Choice of Left		Choice of Right		
	Correct	Incorrect	Correct	Incorrect	
L	19	7	4	19	49
R	7	15	21	4	47

response split. Take, for example, an *S* who consistently alternates left and right choices. From the over-all response split (50:50), the response to events on the immediately preceding trial might also be expected to be of a 50:50 order, with *S* repeating a preceding response half the time and changing half the time. In fact, the contingent responses would be radically different from 50:50, and phi would be at a maximum. Phi will also be high if *S* splits his responses in 50:50 fashion over the total series, but shows a strong tendency to repeat correct choices and switch after incorrect choices. In contrast, phi will be low if *S* often changes his response to a preceding event, e.g., sometimes repeating, sometimes changing after a correct choice. Phi will also be low if *S* consistently chooses only one alternative, so that his responses are not affected at all by success or failure on the previous trial. In all these examples, phi can be said to reflect the extent to which *S*'s responses show a consistent reference to events on the immediately preceding trial.

EXPERIMENT I: SLEEP LOSS

Method

Procedure.—In the baseline period, each of 17 *S*s was tested once a day for 5 consecutive days. Ten of these *S*s were then deprived of sleep for 72 hr., with testing after 28, 52 and 66 hr. of sleep loss. The seven control *S*s were tested at the 28- and 52-hr. points, but not at the 66-hr. point which occurred in the early hours of the morning. All *S*s were given one test in the recovery period, either three or four days after the end of sleep loss. The two-choice task was one of a large number of tests, and the general conditions and setting for the experiment are described in Williams, Lubin, and Goodnow (1959).

The *S*s were told that the tape which programmed the correct choices would be changed from day to day but would not be radically different from previous days.

Schedules.—There were five different schedules, each 144 trials in length. To keep these schedules comparable, all conditional probabilities were controlled. The latter are known to affect the manner in which *S* responds to the outcome of a previous choice (Goodnow, 1955; Hake & Hyman, 1953). They were controlled by determining the expected number of runs of each length in a 50:50 series (Bush & Mosteller, 1955, pp. 214–215). These runs were then strung together in random fashion with one restriction: an equal or close to equal division of stimulus events had to occur by the end of each block of 48 trials.

Tapes 1, 2, 3, 4, and 5 were used consecutively on the five baseline days; Tapes 2, 3, and 4 during the sleep loss period, and, to allow maximum comparability, Tape 4 was repeated again on the recovery day. For all schedules, the sequential analysis covers Trials 47–144.

Subjects.—The 17 *S*s were Army enlisted men, volunteers from a pool of conscientious objectors who offer themselves as *S*s in medical research projects.

Results

Changes in the importance of the just previous trial are shown in Fig. 1, which presents the changes in the critical measure (phi). For experimental *S*s there is a marked rise in phi after 24 hr. of sleep loss and a drop to the baseline level upon recovery. In contrast, control *S*s show a decline in phi from the baseline to the sleep loss period, and a rise in phi at recovery testing. The data in Fig. 1 are presented in terms of change from a baseline score, since the control and experimental *S*s did not turn out to be well matched for phi. (The mean phi scores for the first baseline day are .27 for experimentals and .39 for controls.)

Beyond the first 24 hr., increasing hours of sleep loss do not produce any

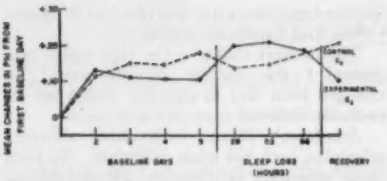

FIG. 1. Effect of sleep loss on phi scores (phi reflects the effect of local events on a choice).

marked increase in mean phi scores. For experimental *Ss*, the average rank-order correlation coefficient with hours of sleep loss is .42 (Lyerly, 1952). What happens is that all sleep deprived *Ss* increase their phi score on one or more of the three sleep loss testings, but *Ss* vary in the day on which their highest score occurs. For this reason, we have averaged the baseline and recovery phi's and compared this score with the mean average phi during the sleep loss period. The value of *t* from this comparison for experimental *Ss* is 2.46, significant at the .01 level on a one-tailed test.

A low phi score occurs under two conditions: (*a*) when *S* always chooses the same alternative, and (*b*) when *S* has a variety of ways of responding to the event on the preceding trial. The latter condition is the important one in sleep loss (the first one does not occur at all).

Before sleep loss, *S* tries to follow and anticipate the vagaries of the stimulus series. Some of the wilder anticipations drop out after the first day's experience, leaving *S* with a preferred "system" which he varies to fit with the trends he sees in past events. During sleep loss, these variations occur less and less often and *S* adopts a simple general rule—repeating a successful choice until he loses, for example. The *S* regards his use of a single rule of thumb as a poorer grade of performance. With this relative failure in mind and with recovery from sleep loss, he returns to his earlier attempts to "outsmart" the machine by anticipating its performance.

Similar changes in the effort put into keeping track and anticipating appear to underlie the controls' performance. These *Ss* placed little value on the role of control, a role they were assigned at the onset of the baseline period (Williams, et al., 1959). They showed their greatest interest during the sleep loss period when they felt themselves to be in direct competition with the experimental *Ss*; this is the only time when their phi scores were lower than those of experimental *Ss*. Recovery testing, regarded by most control *Ss* as meaningless, was the time of lowest interest and highest phi scores, with all but one control *S* giving a higher phi score than on the previous testing.

EXPERIMENT II: ELECTROCONVULSIVE SHOCK (ECS)

Method

Procedure.—On the same machine as in Exp. I, experimental *Ss* were tested before treatment, after 10 ECS, after 20 ECSs, and again one week after the end of treatment. The intervals between the four testings were approximately 1, 3, 3, and 1 week(s). Control *Ss* were tested with the same intervals, but received no ECS. For experimental *Ss*, there was the additional restriction that testing should be 24–48 hr. after an ECS, to avoid some of the more immediate and transient effects of a shock.

Subjects.—The experimental *Ss* were drawn from patients on an Army NP service, for whom ECS had been advised. The *Ss* were all schizophrenic and were recommended by psychiatrists on the NP service as being able to cooperate in testing. In all, 20 *Ss* were given the first test; 8 *Ss* completed the four testings, with the remaining *Ss* becoming unavailable for a variety of administrative reasons (recommendation for ECS changed for 4 *Ss* before treatment started, for 4 *Ss* after 10 ECSs; 4 *Ss* placed on 30 days' leave before recovery testing).

The control *Ss* were nonschizophrenic, since it was difficult to obtain schizophrenic *Ss* who were not receiving medication with unknown effects on two-choice performance. The control *Ss* selected were patients on an amputee ward, in most cases on the ward for prosthetic treatment. These *Ss* were the most likely to be available for the 8-week

period that the experiment covered: they were also in general good health. Of the 16 initial *Ss*, 6 completed all four testings, the remainder being either on leave at the time of one testing, or else either discharged or transferred before the end of testing.

Schedules.—The same 67:33 schedule was used for all four testings. The schedule covered 120 trials, and the series was developed by the same method of combining runs that was described in Exp. I. The 67:33 requirement was approximated for the two halves of the series (41 lefts in the first 60 trials, 39 in the second 60 trials). The sequential analysis covered Trials 19–120.

Results

All results are based on *Ss* who completed all four testings. The two groups of *Ss* start from the same baseline point, but experimental *Ss* show a marked and progressive increase with ECS treatments. (The changes over Tests 1, 2, and 3 yield an average rank order coefficient of + .75 for experimentals, as against − .33 for controls.) The *Ss* given ECS also show a marked mean drop in phi after the end of treatment, but the variance is considerable. (A comparison for experimental *Ss* of scores after 20 ECSs with recovery scores yields a t value of 1.90, significant at the 5% level on a one-tailed test.) The drop is nonetheless encouraging, both because the sample is small and because recovery after ECS is known to be variable.

With the sleep loss *Ss*, changes in phi reflected an increasing use of a rule of thumb which could be described as pivoting around one or at most a few preceding events. Much the same effect occurs with the *Ss* given ECS, although these *Ss* vary considerably more among and within themselves in the procedures that they use. One *S*, for instance, started off with a tendency to alternate responses (73% of his responses were alternations). After 10 and 20 ECSs, respectively, the percentage of alternating responses rises to 92% and 93% and drops on recovery testing to 78%. A second *S* started with a tendency to repeat correct choices. After 10 ECSs, this strategy became more pronounced but after 20 ECSs it gave way to very marked response alternation (88% of responses). A week after the end of ECS, this *S* adopted a mixture of his two previous strategies. Where both of these *Ss* showed response alternation as a dominant pattern, however, a third abandoned an initial alternation and gave his responses in long series of choices. After spending the first 24 trials in short-run choices of left and right, he gave 33 choices of right, followed by 34 of left, and 29 of right again. (The right-hand key, incidentally, pays off 33% of the time.) After the end of ECS, this *S* went back to a mild display of alternation.

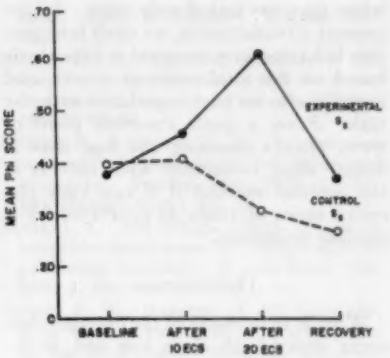

FIG. 2. Effect of electroconvulsive shock (ECS) on phi scores.

In one respect, however, some *Ss* in Exp. II displayed behavior which did not occur at all in Exp. I. This consisted of always choosing the key which was correct 67% of the time. Two experimental *Ss* showed this behavior only upon the fourth (recovery) testing; two control *Ss* approached it on the second testing, and, by the fourth testing, four of the six control *Ss* were giving a heavy majority of choices to the 67% key (85%, 86%, 91%, and 100%, respectively). This is unusual behavior for human *Ss* with a 67:33 task, at least

when they are tested only once. In the present circumstances, we shall interpret this behavior as a reasonable hypothesis based on the total series of events, and possibly also on past experience with the task. From a game theoretic point of view, always choosing the key that is linked most frequently with success is the optimal method if *S* can view the entire series of trials as generated by a random procedure.

Discussion

What is to be made of the changes that occur under both sleep loss and ECS? As long as interest is limited to a statistical statement of the change, there is no problem. Phi is an extremely sensitive index of change, and it rises and falls in a very satisfactory fashion. What does the change mean, however, in terms of *S*'s problem solving behavior?

To answer this question, consider the common aspect of the several types of behavior which can give rise to a low value of phi. In general, the common aspect is the attention given to a long series of stimulus events. This attention most often leads *S* toward varying his strategies and hypotheses in order to take account of the way "things have been going," of a trend perceived in past stimulus events, or the poor rewards of a particular procedure. Less often it leads *S* to give all his choices to the alternative which pays off more frequently. With our *S*s, such 100:0 behavior has several features. It appears only when one alternative does pay off more often than the other; it does not occur on *S*'s first experience with the task; with *S*s given ECS it appears only on recovery testing; and it is based, according to *S*s, on all their past experience with the task. Because of these features, both 100:0 behavior and variations in hypothesis appear to be indicative of attention to a long series of past events.

In contrast, a high value of phi goes with attention to a shorter series of past events. Out of some smaller part of his past experience, *S* comes to adopt a single strategy or procedure and to pursue it without regard to the times when his procedure does not work well or, in some cases, without regard to the over-all characteristics of a stimulus situation. An *S*, for example, who follows the procedure of alternating responses in a 67:33 situation is ignoring completely the frequency with which the stimuli occur. It is as if he cannot set his sights far enough to see what the relative frequency is.

It must be clear that this interpretation of phi as an index of attention to a long series of events is not without some weaknesses. When, for example, *S* gives all his choices to one alternative, other criteria are needed to determine whether this 100:0 behavior stems from a rational, "all things considered" type of approach or whether it reflects the unimaginative behavior that a paradise fish exhibits (Bush & Wilson, 1956).[2] In a similar way, when phi is low because *S* has responded in a variety of ways to past events, other evidence is needed to determine whether *S* is modifying his procedure in a way that fits with past events or simply varying procedures wildly and randomly. Actually, frequent and random shifts in strategy appear to be more a statistical than a psychological possibility, a point which has been made with particular force in the extensive experiments by Ross and Levy (1958).

Despite these difficulties—and few summary measures are without them—there are considerable advantages to a measure which disregards the specific form of *S*'s behavior and asks only about the range of events on which his procedure is based.

It is also worth noting that the same general concept can cover the changes observed with sleep loss and with ECS in the present study, the differences found by Miller and Frick (1949) between children and a rat on a four-choice task, by Leary, Harlow, Settlage, and Greenwood (1952), between normal and

[2] For this reason, we would recommend the use of a 50:50 task where the likelihood of 100:0 behavior is minimal.

brain-damaged monkeys on a double-alternation problem, and by Hyman and Jenkin (1956) between human *Ss* who were told that a two-choice problem was soluble and *Ss* who were told that it was not. In these examples, and also in Frick's (1953) study of anxiety, the particular way of measuring the effect of past events varies in keeping with the several tasks used, but the over-all result is the same: conditions that lower ability or motivation upset the usual relationships between a present response and past events, and the upset is measurable.

Summary

In two experiments, sleep loss and electroconvulsive shock were found to produce a common effect: a restriction in the range of past events upon which *S* bases a response in a two-choice task. The results underline the usefulness of sequential analysis of responses for measuring changes in problem solving behavior.

References

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Bush, R. R., & Wilson, T. R. Two-choice behavior of paradise fish. *J. exp. Psychol.*, 1956, **51**, 315–322.

Frick, F. C. The effect of anxiety—a problem in measurement. *J. comp. physiol. Psychol.*, 1953, **46**, 12–123.

Goodnow, J. J. Determinants of choice-distribution in two-choice probability situations. *Amer. J. Psychol.*, 1955, **68**, 106–116.

Grant, D. A. The discrimination of sequences in stimulus events and the transmission of information. *Amer. Psychologist*, 1954, **9**, 62–68.

Hyman, R., & Jenkin, N. W. Involvement and set as determinants of behavioral stereotypy. *Psychol. Rep.*, 1956, **2**, 131–146, Monogr. Suppl. No. 3.

Janis, I. L. Psychologic effects of electric convulsive treatments: II. Changes in word association reactions. *J. nerv. ment. Dis.*, 1950, **3**, 383–397.

Leary, R. W., Harlow, H. F., Settlage, P. H., & Greenwood, D. D. Performance on double-alternation problems by normal and brain-injured monkeys. *J. comp. physiol. Psychol.*, 1952, **45**, 576–584.

Lyerly, S. B. The average Spearman rank correlation coefficient. *Psychometrika*, 1952, **17**, 421–428.

Miller, G. A., & Frick, F. C. Statistical behavioristics and sequences of responses. *Psychol. Rev.*, 1949, **56**, 311–324.

Ross, B. M., & Levy, N. Patterned predictions of chance events by children and adults. *Psychol. Rep.*, 1958, **4**, 87–124, Monogr. Suppl. No. 1.

Welford, A. T. The psychologist's problem in measuring fatigue. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue.* London: Lewis, 1953. Pp. 183–192.

Williams, H. L., Lubin, A., & Goodnow, J. J. Impaired performance with acute sleep loss. *Psychol. Monogr.*, 1959 (in press).

(Received January 9, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

EFFECTS OF ROUND NUMBER PREFERENCES IN A GUESSING TASK

BRUCE M. ROSS

Rutgers University

AND TRYGG ENGEN

Brown University

The human observer tends to show highly overlearned preferences for certain numbers or terminal digits in the decimal system. In particular, the tendency to favor numbers ending in 0 and 5 (both called "round" in the present context) is commonly observed and has been given occasional passing mention in the literature of experimental psychology (e.g., Barber & Garner, 1951; Chapanis, Garner, & Morgan, 1949; E. P. Reese, T. W. Reese, Volkmann, & Corbin, 1953; Stevens, 1956; Yule, 1927). The thesis of the present study is that round number preferences can be demonstrated by perseverative measures in a guessing task as well as by naive choice. The hypotheses tested were that (*a*) guessing numbers ending in 0 and 5 would be significantly easier than guessing numbers ending in 2 and 7 and that (*b*) subsequent guessing of numbers ending in other digits (i.e., 1, 3, 4, 6, 8, and 9) would be more difficult following prior guessing of numbers ending in 0 and 5 than following numbers ending in 2 or 7. That is, initial success with round numbers will increase the trials required to guess nonround numbers, because immediate confirmation of the round number tendency will encourage more perseverative guesses than success with nonround numbers. This paper presents evidence supporting both hypotheses.

METHOD

Subjects.—The *S*s were 62 women students from Brown University and the University of Rhode Island. An attempt was made to select homogeneous groups of *S*s who had no special number facility.

Procedure.—The instructions, which were the same for all groups, will explain the guessing task. The *E* read the following to an individual *S*:

"This experiment is concerned with your skill in making predictions under conditions of minimum information. We are particularly interested in assessing your skill in comparison to that of the average college student. On the back of each of these cards is written a number between 1 and 120. Some of the numbers may be alike or they may all be different; they have been selected by chance. I will tell you only three things during the course of your prediction. First, if your prediction is within 10 numbers greater or less than the number on the card, I will say, 'Within 10.' Second, if your prediction is more than 10 numbers away from the number on the card but 20 or fewer numbers away, I will say, 'Within 20.' Third, if your number is simply more than 20 numbers away either way, I will say, 'No.'

"Of course, when you make a correct prediction I will tell you. Then you should see if you can use fewer predictions in making a correct prediction of the number on the next card. The fewer the number of predictions it takes you, the better your performance and the higher your score. We are particularly interested in your over-all number of predictions.

"Let me give a couple of examples. If the number on the card were 68, I would say 'Within 10' if the number you predicted was any number between 58 and 78, including the numbers 58 and 78 themselves; I would say 'Within 20' if the number you predicted was between 48 and 88, inclusive. (Sample card shown.) Let me ask you if the number were 31, what would 'Within 10' mean? 'Within

20?' That's right, simply add and subtract 10 and add and subtract 20. The number you get by adding or subtracting 10 or 20 is the end number in the 10 or 20 range, depending on which one you are figuring. Are there any questions?"

For the control group 18 cards were spread face down on a table. The correct number and the "within 10" and "within 20" number ranges were indicated on the card. The seated *S* pointed out a single card which was picked up and held by *E*. The cards were presented to Groups I and II in a predetermined order. The *S* made oral guesses which were recorded by *E*, and she was shown the number on the card after she had guessed it.

In the control group, 20 different series of 18 numbers were selected so that every number from 1 to 120 was guessed three times and for any individual number series only three numbers were drawn from any consecutive 20 numbers, i.e., 1–20, 21–40, etc. Two guessing series were replicated for a total of 22 *Ss*.

In Group I, numbers ending in 2 and 7 were guessed for the first nine numbers and numbers ending in the other digits for the last nine. The number order administered to 20 *Ss* was 17, 72, 22, 92, 77, 112, 17, 57, 42, 54, 115, 21, 43, 9, 98, 70, 86, and 29.

In Group II, the numbers to be guessed were similar to those for Group I. The first nine numbers were lowered by two to convert them to round numbers ending in 0 or 5. The last nine numbers were the same as for Group I except the numbers ending in 0 or 5 (115 and 70) were raised by two so that no round number would reoccur. Thus, the number order administered to the 20 *Ss* in Group II was 15, 70, 20, 90, 75, 110, 15, 55, 40, 54, 117, 21, 43, 9, 98, 72, 86, and 29. Numbers for both Groups I and II used only two terminal digits (2 and 7, 0 and 5, respectively) for the first nine numbers and every other digit not previously represented plus one repetition of the terminal digit 9 for the remaining nine numbers.

Results and Discussion

In the control group, the magnitude of the correct number did not significantly affect the number of guesses per number. This was shown by analyzing the number of guesses made in consecutive series of 20 in the range 1–20, 21–40, etc. The mean number of guesses ranged from 9.42 per number in the 61–80 range to 10.98 in the 101–120 range with an over-all mean of 9.92. Practice effects were also inconsequential as the mean guess per number for the last half of the number series was .18 of a guess more than the first half.

There was a definite round number favoritism. For all guesses, 32% ended in 0 and 13% in 5, while for any one of the other eight digit endings the range was between 5% and 8% of the total. If the first three guesses (or fewer if the number was guessed sooner) for each number are totaled separately, favoritism is much more pronounced—74% of the guesses end in 0, 13% in 5, and 13% in one of the other eight digits. Obviously *Ss* preferred to guess grossly by tens before switching to other digit endings.

Since *Ss* customarily use round numbers most frequently, one would expect round numbers to be guessed in fewer attempts. This proved to be the case. The mean number of guesses was 5.25 for numbers ending in 0; 8.25 for numbers ending in 5; and for other digit endings the range was from 8.68 for numbers ending in 2 to 12.97 for those ending in 3. Two-tailed t tests indicated that the mean difference in *Ss*' performance on numbers ending in 0 and 5 was significant ($P < .001$), but the mean difference between numbers ending in 3 and 5 was not ($.05 < P < .10$). On all counts, thereiore, it can be concluded that numbers ending in 5 are appreciably less "round" than those ending in 0.

The over-all results for Groups I and II are graphed in Fig. 1 in which total mean guesses to correct numbers, including the guess of the correct number itself, is plotted as a function of the particular number guessed. Figure 1 shows a confirmation of the

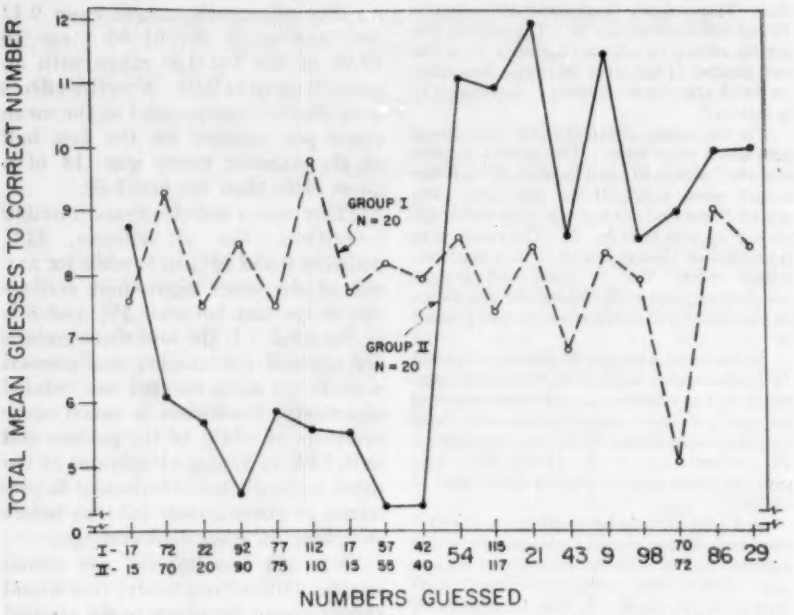

FIG. 1. Total mean guesses to guess correctly two different but closely related number series.

previously stated hypotheses. Because of positive skewness in the raw score distributions found when plotting medians, a square-root transformation was made before between-group significance tests were computed. A t test computed using total guesses per S for the first nine guessed numbers confirmed the first hypothesis; significantly fewer guesses were taken to attain round than nonround numbers plus two higher.

It should be noted that Ss in Group II perform about equally well on numbers ending in either 0 or 5. In comparing results from Group II with those obtained for the control group it appears that 5 becomes "rounder" in the context of being associated with numbers ending in 0. Medians obtained for Group II underscore this result, as the median is six for the first guessed number but four for the next eight numbers ending in 5 or 0. The five-ending number, 15, was not guessed as quickly as 17 when first, but in the context of 5 and 0 endings was guessed in significantly fewer guesses than 17 when in seventh place ($P < .05$).

The important comparison for confirmation of the second hypothesis is between the last nine guessed numbers of Groups I and II in Fig. 1. Group II, which gave significantly fewer guesses on the first nine numbers, gave significantly more on the last nine as shown by a total guess, between-group t test comparison ($P < .001$). Presumably the occurrence of the round numbers 115 and 70 among the last nine numbers of Group I, admitted to prevent the reoccurrence of numbers ending in 2 and 7, could be the major factor in producing a significant group difference. Accordingly,

a t test was also performed on individual total guesses on the seven remaining nonround numbers that were the same for both groups. The round number contribution was not crucial in obtaining a significant t value ($P < .01$).

A specific analysis for perseverative effects was performed by counting the number of guesses per number ending in 0 or 5, excluding correct guesses. The results of this tally are graphed in Fig. 2 where mean guesses ending in 0 or 5 are plotted against the number guessed. For the first nine numbers guessed there is no appreciable difference between Groups I and II in the number of round number guesses used. The round number tendency is operative for Group I despite no successes; the difference between groups noted for the first half of the number series in Fig. 1 is brought about by the necessity for Group I to guess additional nonround numbers in order to achieve correctness.

An analysis of variance of total mean guesses for the last nine guessed numbers of the four-fold guess classification in Fig. 2 showed a significant between-group difference at the .01 level ($F = 35.50$ with 3 and 76 df). Computation of the second-half analysis of variance without the 115–117 and 70–72 comparisons, performed because of the possibility of inflating differences by including guessed round numbers, also proved significant at the .01 level ($F = 28.78$ with 3 and 76 df). The major t test comparisons for total guesses with the last nine numbers further substantiated the perseverative guessing hypothesis; Group II used significantly more zero-ending guesses and significantly more five-ending guesses

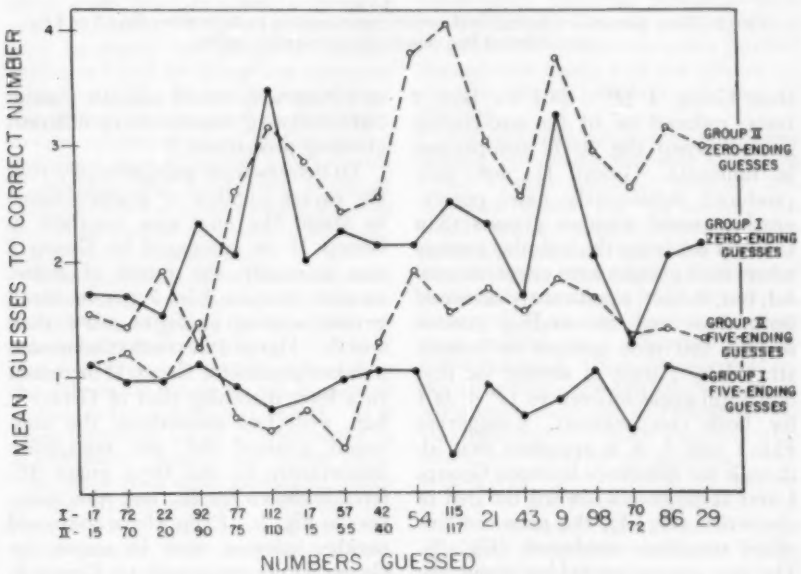
MEAN GUESSES TO CORRECT NUMBER
4
3
2
1
GROUP II ZERO-ENDING GUESSES
GROUP I ZERO-ENDING GUESSES
GROUP II FIVE-ENDING GUESSES
GROUP I FIVE-ENDING GUESSES
I - 17 72 22 92 77 112 17 57 42 54 115 21 43 9 98 70 86 29
II - 15 70 20 90 75 110 15 55 40 117 72
NUMBERS GUESSED

FIG. 2. Mean guesses (excluding final correct guess) ending in 0 or 5 for two different but closely related number series.

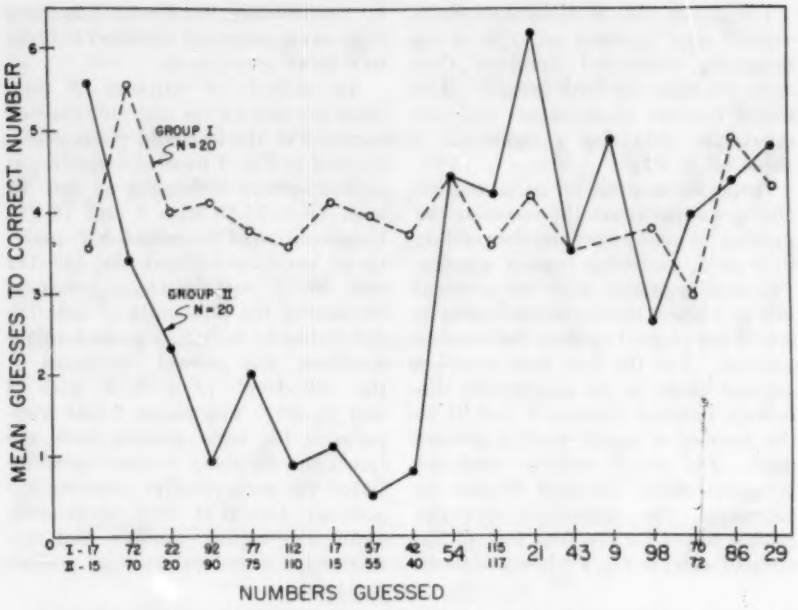

FIG. 3. Mean guesses (excluding final correct guess) ending in digits other than 0 or 5 for two different but closely related number series.

than Group I ($P < .001$ for both t tests; reduced to .01 for zero-ending guesses when the 70–72 comparison is omitted). Group II not only produced significantly more perseverative round number guesses than Group I solely for the last nine guesses where such guesses were never successful, but, in fact, significantly *increased* both five- and zero-ending guesses for the last nine guesses as demonstrated by t tests of second vs. first half total-guess differences ($P < .001$ for both comparisons). Comparing Fig. 1 and 2, it is apparent that although the difference between Groups I and II decreases toward the end of the series (Fig. 1), the perseverative effect continues unabated (Fig. 2). The two groups are taking about the same number of guesses at the end of the series, but due to initial success or failure with round number guesses *Ss* in the two groups are using different guessing procedures.

To demonstrate unequivocally that the excess number of guesses taken to attain the final nine numbers in Group II as compared to Group I was primarily the result of round number guesses, Fig. 3 graphs mean guesses ending in digits other than 0 or 5. Figure 3 shows that nonround number guesses for Group II increased to a level matching that of Group I, but, with two exceptions, the nonround guesses did not contribute importantly to the total guess difference shown for the last nine numbers in Fig. 1. Even where nonround number guesses were in excess for Group II as compared to Group I, they but acted to increase group differences calculable from round num-

ber guess tabulation. Like the control group, Groups I and II tended to use round number guesses at first before switching to nonround guesses. Figure 3 can be interpreted as resulting from the continuation of round number guesses beyond a point of efficiency in guessing single numbers. No attempt was made to analyze single guessing series sequentially because of the wide range of guesses. (In Groups I and II 16 numbers were guessed immediately while two numbers were reached in 26 guesses.)

It was the intent of this experiment to demonstrate the round number tendency as a biasing effect within the framework of a problem solving situation rather than cite its importance for perceptual experiments, as is the case with the references listed in the introductory paragraph. The increased use of round number guesses in a problem solving strategy is postulated as a perseverative effect from prior practice in guessing round numbers. This result contrasts with no second half excess for guesses ending in 2 or 7 for Group I as compared with Group II. It can be argued that the present experiment does not give a unique measure of round number preference—nor does extinction as a measure of learning—but the present study at least extends the acknowledged scope of the round number tendency by illustrating potential aftereffects.

Some comparisons between obtained results and efficient strategy solutions are enlightening, although the use of efficient rational strategies—such as a half-split guessing approach—would not explain the differential results for round number guesses between Groups I and II. If *S* made maximum use of all information in a systematic fashion, she would never take more than eight guesses to arrive at the correct number, including the correct guess itself, even though the worst possible luck prevailed. Therefore an *S* guessing systematically should show a mean guess per number well under eight. Examination of individual *S*s' records failed to find any *S* performing with a consistently efficient strategy —any such use would have run counter to the significant perseverative effects previously shown. (It would have required some effort to be efficient since all guesses were made orally, and *S*s did, on occasion, repeat themselves.) As corroborating evidence, group results indicated that a general use of efficient strategies did not develop, since number of guesses did not decline from the initial inefficient levels. Although an exact solution has not been worked out, guessing of numbers ending in 0 and 5 in a median number of four guesses, performed by *S*s in Group II on the second through ninth numbers, is superior to the best strategy solution which could be evolved from information given in the preliminary instructions. Indeed, if *S* had been told that the number to be guessed was one of the 24 numbers between 5 and 120 ending in 0 or 5, no strategy solution could on the average bring the median number of guesses below four! Since *S*s performed with great efficiency only in guessing round numbers, the conclusion seems justified that number habits were not eclipsed by possible efficient strategy solutions of the task.

The results of the present experiment confirm the observation that there is a round number tendency; however, this tendency should not be invoked too readily without direct measurement since numerous context effects are possible. For example, Hartmann (1929) found that numbers ending in 0 and 5 embedded in lists of 25 one- or two-digit numbers were recalled at a frequency slightly below the mean. Best recalled were one-digit numbers among two-digit numbers, while repetitions such as 22 and 33 were also well recalled. These results demonstrate the relativity of "roundness" resulting from context effects, as does the finding in the present experiment that numbers ending in 5 are "rounder" when neighbors of numbers ending in 0 than in more mixed company.

Results of this study have direct cautionary bearing on psychophysical

procedures wherever an interaction may exist between number habits and more purely judgmental processes. Further, this experiment outlines a method of describing by perseverative effects the strength of a nonlaboratory overlearned response bias. Finally, the substantiation of any consistent number habit for the cardinal integers can scarcely be trivial since the decimal number system is more universal than all other languages.

Summary

Three groups of college women (N = 22, 20, and 20) individually guessed series of 18 numbers in the 1 to 120 range. They were informed by E only if their guess was within 10 or 20 numbers either way from the number designated correct. Guesses continued until the correct number was reached. The control group showed marked favoritism for guesses ending in 0 and some preference for guesses ending in 5, but magnitude of the number to be guessed and practice did not influence results.

The main comparison was made between a group which had correct numbers ending in 0 or 5 ("round" numbers) for the first half of the guessed series and a group which guessed numbers ending in 2 or 7. Both groups were switched for the second half of their number guessing series to "nonround" correct numbers. As predicted, the group guessing numbers ending in 0 and 5 took significantly fewer guesses than the group guessing numbers ending in 2 and 7 and significantly more guesses when the switch was made to guessing nonround numbers. When number guesses ending in 0 and 5 were tallied, it was found that round number guesses were initially as frequent for both groups, but that round number guesses significantly increased for the group which had beginning success predicting round numbers after these numbers ceased to be correct. Perseverative effects from success with round number guesses continued longer than total guess scores indicated.

It was pointed out that round number effects, though strongly structured by the decimal number system, are susceptible to numerous context effects. For example, a number ending in 5 is only slightly "rounder" than other numbers, if its appears nonsystematically, but it is a consistent round number if coupled with numbers ending in 0. Zero-ending numbers were "round" in all experimental contexts.

References

Barber, J. L., & Garner, W. R. The effect of scale numbering on scale-reading accuracy and speed. *J. exp. Psychol.*, 1951, **41**, 298–309.

Chapanis, A., Garner, W. R., & Morgan, C. T. *Applied experimental psychology.* New York: Wiley, 1949.

Hartmann, G. W. Determination of easy and difficult number combinations by immediate auditory memory. In Proc. Ninth int. Congr. Psychol., Princeton, 1930.

Reese, E. P., Reese, T. W., Volkmann, J., & Corbin, H. H. (Eds.) *Psychophysical research summary report 1946–1952.* NAVEXOS P.-1104, Tech. rep. SPECDEVECEN 131-1-5., January, 1953.

Stevens, S. S. The direct estimation of sensory magnitudes—loudness. *Amer. J. Psychol.*, 1956, **69**, 1–25.

Yule, G. U. On reading a scale. *J. Roy. Statist. Soc.*, 1927, **90**, 570–587.

(Received February 13, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

SENSORY PRECONDITIONING: CENTRAL LINKAGE OR RESPONSE MEDIATION?[1]

S. H. LOVIBOND

University of Adelaide

Despite the methodological inadequacies of many of the earlier studies of sensory preconditioning (Reid, 1952), the existence of this phenomenon now seems to be generally accepted. Accordingly, recent experiments in this field have been directed towards elucidation of the mechanism and significant parameters of this form of learning.

The sensory preconditioning (SPC) experiments have been interpreted by some S-S theorists (Birch & Bitterman, 1951) as offering clear-cut support for their position. However, several writers (Coppock, 1958; Silver & Meyer, 1953; Wickens & Briggs, 1951) have offered an analysis of SPC in S-R terms. According to such an analysis, the transfer observed in SPC experiments takes place on the basis of a mediating, stimulus-producing response or responses. Usually the nature of the mediating response has been left unspecified and it appears to have been assumed that the range of responses which could possibly fulfill the mediating function is very wide.

However, an examination of the role of the mediating stimulus suggests that the characteristics required are such as to limit rather sharply the possible classes of mediating responses.

The first characteristic required of a mediating response is that it should be elicited with a high degree of reliability in the presence of its stimulus. This is necessary if the postulated S-R connections are to be formed in as few as 10 trials (Bitterman, Reed, & Kubala, 1953; Coppock, 1958).

Secondly, the mediating response (and hence the stimulus it produces) must be highly specific to the stimulus giving rise to it. Consider, for example, an experiment such as that of Bitterman et al. (1953), in which two lights, L_1 and L_2, are used as preconditioning stimuli; L_2 is the CS in the training phase and L_1 is the test stimulus in the test phase. If the responses to the two lights are not stimulus-specific but are highly general (in the sense of being elicited by a wide range of stimuli), the sequence of events during the experiment will be as shown in the comparison between groups PC_1 and C_1 in Fig. 1.

In the experimental group (PC_1), transfer occurs in the test phase via the mediator $R_{s2}s_2$ which is a CR of the test stimulus. In the control group (C_1), the test stimulus L_1 gives rise to R_{1s1}, which is little different from $R_{s2}s_2$. If R_{1s1} is elicited strongly as a UCR or much practiced CR of the test stimulus, R_2 should also be elicited strongly as a function of generalization from $R_{s2}s_2$.

Mediation through R_{1s1} is also present in the experimental group, but it is unlikely to add greatly to the strength of R_2 which is already elicited strongly by $R_{s2}s_2$. Any direct stimulus generalization from L_2 to L_1 would be an additional factor tending to equate the responses of experimental and control groups at a high level.

However, given highly stimulus-specific responses to the two lights, the difference between experimental and control groups, as shown in the comparison between PC_2 and C_2 in Fig. 1, should be considerable. In this case the test light, when presented to the experimental group, gives rise to a specific response-produced stimulus, S_2, which has been paired with shock in the training stage. Hence R_2 should be elicited strongly.

On the other hand, in the case of the control group the test light gives rise to a response-produced stimulus, S_1, which has not been associated with shock, and which is distinctively different from the shock-associated stimulus S_2. Hence R_2 should not be elicited except as a result of direct stimulus generalization from L_2 to L_1.

In the light of the foregoing analysis, it

[1] The author is indebted to L. B. Brown for a number of helpful suggestions concerning presentation.

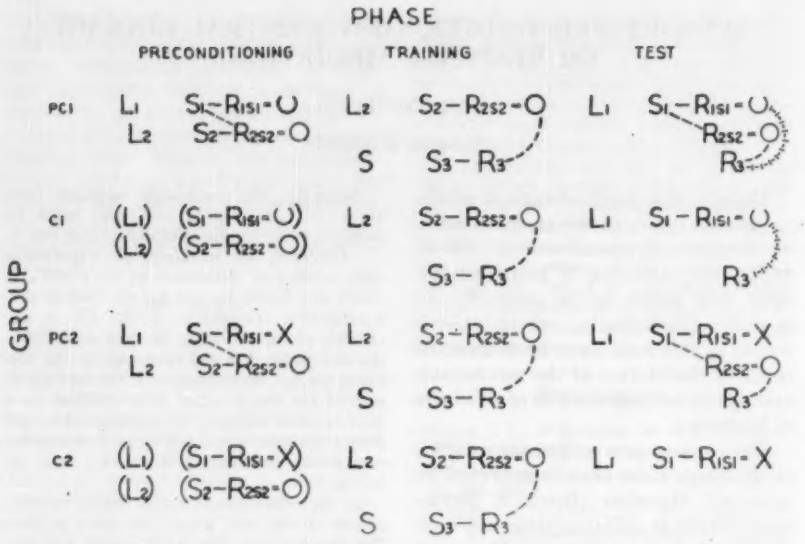

FIG. 1. Showing requirement of stimulus-specificity in S-R mediator. L, light; S, shock; O, stimuli of general responses; X O, stimuli of specific responses; —, unconditioned bond; – – – –, conditioned bond; |–|–|–, generalized bond; ( ) nonpaired stimulus. Displacement of paired stimuli on horizontal axis indicates delay. (Following Coppock [1958].)

would appear that a response as lacking in stimulus specificity as pupillary contraction to a light stimulus, which has been suggested by way of illustration by Coppock (1958), would be quite unlikely to meet the requirements of a mediator.

The type of implicit response which appears most likely to meet these requirements, in the case of human *Ss*, is the subvocal response, particularly the response of naming the stimulus or stimulus compound. Both the specificity and reliability of this type of response should be high. On the other hand it is highly unlikely that nonverbal forms of implicit response (postural, autonomic, and glandular) could embody the tremendous diversity of stimulus-specific responses which would be required by the S-R mediation hypothesis.

Wickens and Briggs (1951), in seeking experimental support for an S-R analysis of SPC, have tacitly recognized that the most plausible hypothesis concerning the nature of the mediating response in human beings is that it is subvocal in character. These authors present evidence which suggests that, under some circumstances at least, SPC transfer may be mediated by stimulus-producing responses. However, it is conceivable that although response mediation may be a *possible* mechanism of SPC transfer, such mediation is not an essential condition for this transfer.

This paper reports an experiment designed to obtain evidence for or against the hypothesis that a *subvocal* mediating response is essential for transfer by comparing the transfer obtained under two sets of conditions: (*a*) conditions favorable for both response mediation and the formation of direct S-S linkages; and (*b*) conditions unfavorable for mediation via a subvocal response, but favorable for the formation of direct S-S linkages.

The first set of conditions was provided by requiring one experimental group (Group + ve VR) to make consistent explicit verbal responses (naming the stimulus) to the preconditioning stimuli throughout all

phases of the experiment. It was assumed that making a consistent overt vocal response in this fashion would strengthen any tendency for mediated transfer to occur via a subvocal naming response, but would not interfere with any direct linkage transfer.

The second set of conditions was provided by requiring another experimental group (Group − ve VR) to make one type of explicit verbal response to the preconditioning stimuli during the preconditioning phase, and different explicit verbal responses to these stimuli during subsequent phases. It was assumed that making an explicit verbal response would preclude the possibility of making a subvocal response of a different character at the same time, without affecting possible direct linkage formation.

The subvocal mediating response hypothesis and the direct linkage hypothesis call for different predictions concerning the performance of the experimental groups.

On the subvocal mediating response hypothesis, transfer should be significantly poorer under the second set of conditions (Group − ve VR). On the other hand, the direct linkage hypothesis predicts no difference in transfer between the two groups.

In addition to a normal or conventional control group in which the preconditioning stimuli were presented separately during the preconditioning phase, an additional group underwent a special control procedure. In this group the test stimulus-to-be was linked in the preconditioning phase with a stimulus which was to become the differentiating stimulus during the training phase. It was hoped by this procedure to link the test stimulus with a stimulus which was to become inhibitory for the test response (GSR) and thus to limit the test phase response of this group to a level below that of the normal control group.

Method

Subjects

The *S*s were 48 adult volunteers, 24 men and 24 women. Most *S*s were students of the Adelaide Teachers' College or the University of Adelaide.

Apparatus

The circuit used to measure the GSR was that described by Davis, Siddons, and Stout (1954), modified to drive a Both single-channel pen recorder. The degree of attenuation employed varied from *S* to *S* between 100 and 200 ohms per mm. of pen deflection. The *S*'s palm was rubbed thoroughly with electrode jelly before attaching the electrodes, which were placed as far apart as practicable on the palm of the left hand.

The 2-ma., 50-cycle AC electric shock which served as the UCS was produced by a constant current electronic stimulator (Lovibond & Turner, 1956) and delivered to electrodes on the tip of the right third finger and the palm of the right hand.

The neutral stimuli used were a red and a white light and a muted buzzer. The light stimuli were provided by two 32-v., 15-w. lamps mounted side by side in a box and arranged to illuminate two circular ground glass apertures 3 in. in diameter and 9 in. apart in the front panel of the box. The buzzer was mounted inside the box and was covered with a layer of felt. All neutral stimuli and the UCS were presented automatically by means of relays operated by a Palmer timer and were recorded by stimulus markers. A switching arrangement permitted selection of any desired sequence of stimuli, and manual operation of a push button was used to start the selected sequence.

Colored slides of outdoor scenes, Rorschach cards, and brilliantly colored moths, designed to offset the development of satiation or inhibition, were presented manually by means of a slide projector.

Two adjacent rooms served as the experimental and control rooms. All apparatus was housed in the control room and a third room nearby. The hum of the recorder was faintly audible in the experimental room and it served to mask other slight sounds from the control room.

Procedure

The *Ss* were divided randomly into two experimental groups, positive verbal reinforcement, (+ ve VR), and negative verbal reinforcement (− ve VR) and two control groups, inhibitory control (IC) and normal control (NC). There were three phases in the experiment for each group.

Preconditioning phase.—The two experimental groups were given 10 paired trials of the buzzer (the test stimulus-to-be) and the stimulus to be conditioned (CS), five trials of the red light alone and five trials of the white light alone in a random sequence which was the same for all *Ss*. Three colored slides were presented at intervals of six stimuli.

For the inhibitory control group the same stimuli were presented in the same sequence, with the exception that the buzzer and the differentiating stimulus (DS) were paired. For the normal control group also the stimulus presentations were the same as those of the experimental groups with the exception that this group received 10 presentations of the buzzer alone in place of the buzzer-CS combination.

The positive verbal reinforcement of the first experimental group consisted in *Ss* responding to the buzzer-CS combination by naming these stimuli aloud, e.g. "Buzzer-Red." The inhibitory control *Ss* responded similarly to the buzzer—DS combination.

The negative verbal reinforcement *Ss* were required to respond with the word "Now" to the buzzer-CS combination, while the normal control *Ss* verbalized "Now" when the CS and the DS were presented separately.

Training phase.—In the second or training phase, all *Ss* received 10 classical conditioning trials of CS-shock-DS. This procedure was adopted in the light of the contention of Konorski (1948) that overlapping a stimulus with the cessation of the UCS provides optimum conditions for the development of inhibitory linkages. Differentiation should thus have been facilitated.

For half of the *Ss* in all groups CS was the red light and DS was the white light. For the remaining *Ss* these stimuli were reversed. Four presentations of CS alone served as conditioning test trials. Ten presentations of DS and three presentations of colored slides were interspersed irregularly with the other trials in an order which was the same for all *Ss*.

Test phase.—In the test phase, all *Ss* received 10 presentations of the buzzer alone and were required to verbalize "Buzzer" as soon as the sound was heard.

Stimulus Time Relations

The duration of each light stimulus and the buzzer was 1 sec. throughout. The duration of the shock was 1.5 sec. In the preconditioning trials with buzzer and light, the onset of the buzzer preceded by .5 sec. the onset of the light. Thus the buzzer and the light were on together for .5 sec. In the reinforced training trials the onset of the CS preceded the onset of the shock by .5 sec., and the onset of the DS preceded the termination of the shock by .5 sec. In all phases the intertrial interval varied randomly between 30 and 45 sec.

Scoring

The unit of GSR measurement used was change in conductance during the 10 sec. following stimulation. For the purposes of analysis five separate scores were computed for each *S*. These were CS_1, the sum of responses to the first two conditioning test trials; CS_2, the sum of responses to the last two conditioning test trials; DS_1, the sum of responses to DS Trials 3 and 4; DS_2, the sum of responses to DS Trials 9 and 10; BT, the sum of the first five buzzer test trial responses.

Results

Considerable positive skewing was evident in the training and test phase scores despite the use of conductance units. Hence nonparametric methods of analysis were used throughout.

Tests of the equivalence of the four groups in terms of GSR reactivity during the training phase were made with the Kruskal-Wallis one-way analysis of variance by ranks (Siegel, 1956). Tests were made on CS_1, CS_2, DS_1, and DS_2 scores. The obtained H values of 1.4, 1.6, 4.5, and 4.2 with 3 df all have probabilities greater than .20. There is thus no evidence to suggest that the groups were drawn from populations different with respect to GSR reactivity.

Figure 2 shows the median response of the combined groups on the conditioning test trials and the median response of each group on the first five buzzer test trials. Each of the first three DS values plotted represents the

mean of the DS trial immediately preceding, and that immediately following, the corresponding CS test trial. The last DS value is the mean of the two DS trials immediately preceding the last CS test trial.

There is a consistent increase in response to CS over the four test trials. Response to DS tends to increase at first and then decreases over the second half of the training session.

In order to test for the occurrence of conditioning and differentiation, a number of tests of paired replicates by ranks (Siegel, 1956) was carried out. Firstly, the significance of the trend towards increased response to CS during the training phase was tested by comparing CS_1 and CS_2 scores in each group. Secondly, in order to test the significance of the divergence between CS and DS scores in the later stages of training, two difference scores were obtained for each *S*. These were the differences between CS_1 and DS_1 and between CS_2 and DS_2. The differences between these difference scores were ranked for each group separately. Negative signs were attached to differences indicating convergence rather than divergence, and the sums of negative ranks were computed. All eight comparisons yielded significant one-tailed probabilities ranging from .005 to .05.

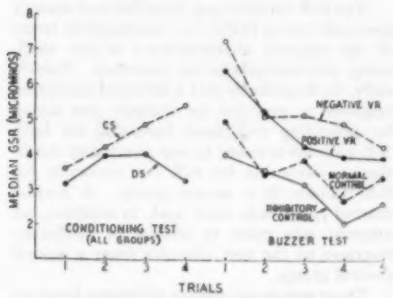

FIG. 2. Median response on conditioning test trials and first five buzzer test trials.

TABLE 1

CORRELATIONS BETWEEN CS_2, DS_2, AND BT SCORES

Group	Paired Scores		
	CS_2-DS_2	CS_2-BT	DS_2-BT
+ve VR	.68	.97	.73
−ve VR	.85	.91	.88
NC	.83	.86	.88
IC	.30	.36	.82

From these results it may be concluded that both conditioning and differentiation occurred. Nevertheless the differentiation achieved was relatively poor as the level of DS_2 responses remained at approximately 70% of the magnitude of CS_2 responses over all four groups.

Figure 2 shows no consistent difference between the two experimental groups on the buzzer test trials, although Group − ve VR tends to give a higher level of response. There is also overlapping between the control groups with Group IC giving a lower median response on four of the five trials. The response levels of both experimental groups are consistently higher than those of both control groups.

As a preliminary check on the effectiveness of preconditioning, Kendall rank-order correlation coefficients were computed between CS_2 and DS_2 scores, between CS_2 and BT scores, and between DS_2 and BT scores. Given the occurrence of preconditioning in both experimental groups, it would be expected that for the experimental groups tau CS_2, BT > tau DS_2, BT; and for Group IC, tau DS_2, BT > tau CS_2, BT, whereas for Group NC, tau CS_2, BT = tau DS_2, BT = tau CS_2, DS_2.

In general these predictions are confirmed as is shown in Table 1. The confirmation is quite striking in the

case of Group IC, but in the other groups the correlations between CS and DS are too high to permit strong confirmation.

The very high correlations obtained, together with large individual differences in GSR reactivity, suggested the possibility of greatly increasing the precision of the test phase score analysis by matching *S*s in terms of CS_2 scores. The differences between the BT scores of matched pairs from the groups contrasted were ranked and negative signs were attached to differences which were in the direction opposite to that of the difference between the medians.

Because this method of evaluation revealed no significant difference between the two experimental groups (sum of negative ranks = 24, $P > .20$), or between the two control groups (sum of negative ranks = 31, $P > .50$), the occurrence of preconditioning was tested by contrasting the pooled BT scores of the two experimental groups and the pooled BT scores of the two control groups. This test showed the difference between the combined experimental and combined control groups to be highly significant (sum of negative ranks = 44, $P < .003$).

Discussion

The nonsignificant difference between the two control groups indicates that the procedure employed with the inhibitory control group was unsuccessful. However, the high DS_2 — BT correlation in Group IC suggests that if more adequate differentiation could be achieved, the IC procedure might be very effective.

The data from the two experimental groups run counter to the subvocal mediating response hypothesis of SPC. On this hypothesis the response level of the — ve VR group should have been lower than that of the + ve VR group because the opportunity to develop a mediating subvocal response to the buzzer was not present in the — ve VR group. The obtained difference, although not significant, is in the opposite direction. This result can be interpreted as evidence that a subvocal mediating response is not essential for SPC transfer.

This interpretation is, of course, based on the assumption that making an explicit verbal response effectively precludes the making of a subvocal response of a different character at the same time. In any case, it is difficult to see how response-mediated transfer could have remained unaffected under the conditions of the — ve VR group.

If a subvocal response is not essential for transfer, it would seem necessary to specify an alternative mediating response in order to maintain the view that response mediation is the *only* mechanism of SPC. However, in view of (*a*) the difficulties of meeting the requirements of stimulus-specificity and reliability of the mediating response; (*b*) previous evidence (Coppock, 1958) against the S-R hypothesis as a complete account of SPC, and (*c*) the evidence of Wickens and Briggs (1951) favoring the mediating response as a *possible* mechanism, it seems more reasonable to conclude that SPC is based primarily on the formation of direct S-S linkages and that stimulus-producing responses provide a possible subsidiary mechanism.

Such a conclusion would be consistent with the findings of many investigators who have studied the relationship between implicit responses and such processes as thinking and silent reading. There seems to be fairly general agreement that implicit responses usually accompany and often contribute to these processes, but such responses are not essential to their occurrence.

Summary

The S-R conditioning hypothesis of sensory preconditioning (SPC) was examined in terms of the required characteristics of the mediating stimulus-producing response. Specifically, the hypothesis that a subvocal mediating response is essential for transfer was tested by providing conditions favorable for both S-R and S-S transfer in one group and conditions unfavorable for S-R but favorable for S-S transfer in a second group. A normal control group was used, and, in addition, an attempt was made to obtain an inhibitory response to the test stimulus from a second control group.

There was no significant difference between the test responses of the two experimental groups or between the test responses of the

two control groups, but the response level of the combined experimental groups was significantly higher than that of the combined control groups.

It was concluded that SPC is based primarily on the formation of direct S-S linkages, with response mediation as a possible subsidiary mechanism.

References

Birch, H. G., & Bitterman, M. E. Sensory integration and cognitive theory. *Psychol. Rev.*, 1951, **58**, 355–361.

Bitterman, M. E., Reed, P. C., & Kubala, A. L. The strength of sensory preconditioning. *J. exp. Psychol.*, 1953, **46**, 178–182.

Coppock, W. J. Pre-extinction in sensory preconditioning. *J. exp. Psychol.*, 1958, **55**, 213–219.

Davis, R. C., Siddons, G. F., & Stout, G. L. Apparatus for recording autonomic states and changes. *Amer. J. Psychol.*, 1954, **67**, 343–352.

Konorski, J. *Conditioned reflexes and neuron organization.* Cambridge, Eng.: Cambridge Univer. Press, 1948.

Lovibond, S. H., & Turner, K. G. A constant current electronic stimulator for use in psychological experiments. *Austral. J. Psychol.*, 1956, **8**, 77–83.

Reid, R. L. A test of sensory pre-conditioning in pigeons. *Quart. J. exp. Psychol.*, 1952, **4**, 49–56.

Siegel, S. *Nonparametric statistics.* New York: McGraw-Hill, 1956.

Silver, C. A., & Meyer, D. R. Temporal factors in sensory pre-conditioning. *J. comp. physiol. Psychol.*, 1953, **47**, 56–59.

Wickens, D. D., & Briggs, G. E. Mediated stimulus generalization as a factor in sensory pre-conditioning. *J. exp. Psychol.*, 1951, **42**, 197–200.

(Received February 5, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

THE EFFECT OF FIXATION CONDITIONS ON DEPTH DISCRIMINATION THRESHOLDS AT SCOTOPIC AND PHOTOPIC ILLUMINANCE LEVELS[1]

ALFRED LIT[2]

Vision Research Laboratory, University of Michigan

The present study is a direct continuation of a previous experiment (Lit, 1959) in which depth discrimination thresholds, based on equidistance (equality) settings of two stimulus rods in a "real-depth" situation, were obtained at various illuminance levels under conditions of equal and increasing amounts of unequal binocular retinal illuminance. In that experiment, the threshold data obtained for the condition of unequal binocular retinal illuminances were essentially new, but the threshold data for equal binocular retinal illuminances were related to data obtained by Mueller and Lloyd (1948) in an earlier experiment on equality settings of targets presented in a stereoscope at various background luminances. A comparison of the results of these two previous experiments on equal binocular retinal illuminances revealed many basic similarities. Significant differences were noted, however, in the over-all range of variation of threshold values and in the specified shape of the threshold-illuminance curves. Accordingly, the present study was designed primarily to determine whether the quantitative differences between the results of the previous experiment and those reported by Mueller and Lloyd could be reconciled by a consideration of the effects of systematic variations in conditions of fixation on depth thresholds.

An auxiliary aim of the present study was to yield new data that could be related to a current controversial issue concerning the relative contributions of stereoscopic parallax and convergence cues to stereoscopic acuity. For example, the experiments of Wright (1951), Rady and Ishak (1955), and Ogle (1956) have obtained depth discrimination thresholds for two given stimulus targets of equal luminance presented under varying amounts of lateral separation. The conditions of fixation typically used in these experiments were restricted to two cases: (*a*) fixation held constant on the stationary standard target, and (*b*) alternating fixation from comparison to standard target. The results reported by each of these investigators showed that, for any given lateral separation of the stimulus targets, the stereoscopic threshold angle was characteristically smaller (that is, stereoscopic acuity was higher) under the condition of fixation involving alternating convergence movements of the eyes than under the condition of constant fixation. When the lateral separation of the stimulus targets was progressively increased, the difference in corresponding threshold values for the two methods of fixation increased. Wright interpreted these results to mean that alternating fixation provides an additional cue to depth based on the accompanying changes in convergence. Ogle, on the other hand, argued that

[1] This work was conducted by Project MICHIGAN under Department of the Army Contract (DA-36-039 SC-52654), administered by the United States Army Signal Corps.

[2] Now with Bendix Systems Division, Ann Arbor, Michigan.

an alternating fixation allows a lower threshold because the images of the stimulus targets fall on retinal regions of higher visual acuity in the two eyes.

The present study investigates the effects of three conditions of fixation on the magnitude of the stereoscopic threshold angle. Two stimulus targets having a fixed lateral separation (0.75°) at 100 cm. are viewed under conditions of equal binocular retinal illuminances ranging from a low scotopic to a high photopic level. The threshold data are based on equidistance (equality) settings of the stimulus targets.

Apparatus and Procedure

A detailed description of the apparatus[2] has been given in the report of earlier experiments (Lit, 1959). This apparatus was originally used in experiments on the Pulfrich stereophenomenon (Lit & Hyman, 1951). It provides a movable comparison target and a stationary standard target that are binocularly viewed by *O* through a pair of circular artificial pupils 2.5 mm. in diameter. The two targets are blackened steel rods ¼ in. in diameter. The comparison rod is located in the lower half of *O*'s visual field. The *O* can adjust the position of the comparison rod along his vertical median plane by turning a wheel located in the dark room in which he is seated. The standard rod is located in the upper half of *O*'s visual field at a fixed distance of 100 cm. from his eyes. The standard rod is laterally displaced 1.3 cm. (0.75°) to the right of *O*'s vertical median plane. The lower end of the standard rod and the upper end of the comparison rod lie in *O*'s horizontal plane of fixation. Uniform background illumination is provided by a lightbox located in a frontal plane 250 cm. from *O*'s eyes. A filter box is provided in front of each eye so that by use of combinations of three neutral-density filters, equal binocular retinal illuminances can be produced at any desired illuminance level. With the 2.5-mm. artificial pupil in use, the retinal illuminance without filters is 4.16 log trolands. Screening units provide *O* with a horizontal rectangular field of view, 21.6° × 4.2°.

Two emmetropic graduate students with normal visual acuity and good binocular functioning served as paid *O*s. The interpupillary separation of Observer F. C. measured 6.20 cm. and that of Observer R. B. measured 6.40 cm. Each was highly trained in making equidistance (equality) settings. In performing this type of depth setting, *O* adjusts the depth position of the comparison target (using a "bracketing" procedure) until the comparison target and the stationary standard target appear to lie in the same frontoparallel plane.

The equality settings were made under three different conditions of fixation. Condition 1 required each *O* to maintain constant fixation on the upper end of the movable comparison rod; the laterally displaced stationary standard rod was never to be viewed directly. In this case, variations in convergence and accommodation necessarily accompanied the movements of the comparison rod that occurred during the bracketing procedure. The movable comparison rod was presumably imaged on each fovea continuously. Under Cond. 2, each *O* maintained steady fixation on the lower end of the laterally displaced, stationary standard rod; the movable comparison rod in the vertical median plane was never to be viewed directly. Under Cond. 3, *O* was allowed to execute free eye movement (that is, alternating fixation) from comparison to standard rod during the bracketing procedure. The rate of fixation alternation was not controlled; each *O* was allowed to fixate at will from comparison to standard rod during the course of his equality setting. In this case, variations in convergence and accommodation accompanied the movements of the comparison rod, but unlike Cond. 1, the movable comparison rod was not continuously imaged on each fovea.

The data for Cond. 1 were obtained during the course of the previous experiment (Lit, 1959). The data for Cond. 2 and 3 were collected immediately following the previous experiment. For all three methods of fixation, equality settings were made under conditions of equal binocular retinal illuminances at 14 selected levels of illuminance, log *E*, ranging from − 2.34 log trolands to 2.75 log trolands. Testing at all 14 values of log *E* under a given method of fixation constituted an experimental session. In a given session, the 14 illuminance levels were presented in order of increasing magnitude. For Cond. 1, four groups of 20 equality set-

[2] The apparatus was originally constructed at the Pupin Laboratories, Columbia University, partially through funds from a research grant-in-aid generously provided by the American Academy of Optometry.

TABLE 1

EFFECT OF CONDITIONS OF FIXATION ON DEPTH DISCRIMINATION (STEREOSCOPIC) THRESHOLDS OBTAINED AT 14 SELECTED LEVELS OF RETINAL ILLUMINANCE, LOG E

log E (trolands)	Cond. 1 Fixation on Comparison Rod				Cond. 2 Fixation on Standard Rod				Cond. 3 Alternating Fixation			
	F. C.		R. B.		F. C.		R. B.		F. C.		R. B.	
*	$\eta_{\Delta R}$	$\eta_{A.D.}$	$\eta_{\Delta R}$	$\eta_{A.D.}$	$\eta_{\Delta R}$	$\eta_{A.D.}$	$\eta_{\Delta R}$	$\eta_{A.D.}$	$\eta_{\Delta R}$	$\eta_{A.D.}$	$\eta_{\Delta R}$	$\eta_{A.D.}$
−2.34	343.6	162.1	−274.3[a]	177.0	542.1	247.8	−371.8	193.6	484.5	218.4	7.9	218.8
−1.94	321.8	101.4	−183.5	114.3	365.9	168.7	−181.9	146.9	372.3	124.3	−155.0	176.4
−1.65	295.1	75.9	−141.1	94.2	312.3	116.6	− 88.7	115.9	355.1	105.9	− 29.9	97.2
−1.55	268.1	70.9	− 89.6	72.8	278.8	109.4	− 67.8	113.4	332.7	94.1	4.7	104.8
−1.25	278.6	66.6	− 70.2	82.8	283.6	80.3	− 66.0	84.0	316.5	76.6	− 17.1	86.9
−0.96	233.8	52.7	− 89.9	58.2	261.6	66.3	− 43.1	59.8	244.3	59.8	− 27.5	67.2
−0.60	136.9	26.5	− 68.3	32.0	162.7	40.8	− 71.4	28.6	182.3	40.4	− 79.3	47.0
−0.26	89.4	18.0	− 41.3	23.1	100.6	27.9	− 41.0	19.7	108.1	20.3	− 51.2	23.7
0.03	62.3	13.2	− 40.4	14.5	69.0	23.4	− 29.7	16.2	74.9	18.3	− 35.1	17.7
0.51	51.9	10.9	− 31.2	11.8	58.8	15.2	− 21.7	14.3	65.3	13.7	− 35.8	11.8
1.09	50.9	10.7	− 25.7	10.3	55.3	12.4	− 20.6	10.2	57.9	11.9	− 32.7	9.4
1.77	53.3	10.8	− 27.5	10.8	54.8	11.2	− 21.0	7.9	61.6	9.8	− 35.3	9.0
2.07	47.3	9.3	− 28.9	8.8	51.0	12.0	− 22.1	7.9	54.4	11.6	− 29.0	8.4
2.75	45.1	8.3	− 28.8	9.4	44.3	9.6	− 22.4	8.4	44.0	7.9	27.1	7.8

Note.—The stereoscopic threshold angle, $\eta_{A.D.}$, is given in seconds of arc for Observers F. C. and R. B., based on the mean value of the average deviations of several sets of 20 equality readings taken under each condition of fixation and retinal illuminance. The corresponding mean angular magnitude of the constant errors of the equality settings is given by $\eta_{\Delta R}$.

[a] A negative value of $\eta_{\Delta R}$ indicates that the mean distance of the comparison rod from O's eyes was less than that of the fixed standard rod.

tings (80 readings) were made by each O at each illuminance level; for Cond. 2 and 3, three groups of 20 equality settings (60 readings for each method) were made by each O at each illuminance level. In the 10 experimental sessions held for each O, the data were obtained after a 30-min. period of dark adaptation. During the course of a given session, a 3-min. period of light adaptation followed each successive increase of illuminance.

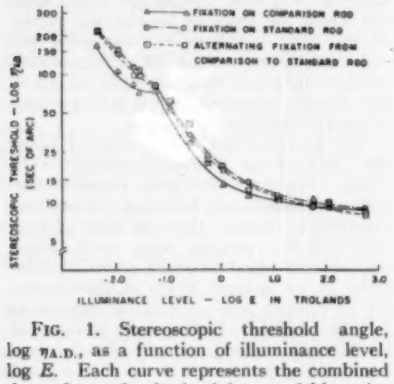

FIG. 1. Stereoscopic threshold angle, log $\eta_{A.D.}$, as a function of illuminance level, log E. Each curve represents the combined data of two Os obtained for equal binocular retinal illuminances under the three conditions of fixation.

RESULTS

Table 1 summarizes the results obtained from both Os. The data are given in terms of the magnitude of the stereoscopic threshold angle ($\eta_{A.D.}$) and the angular magnitude of the constant error ($\eta_{\Delta R}$). Each entry of $\eta_{A.D.}$ is based on the mean value of the average deviations of the respective groups of 20 equality settings obtained at each illuminance level under the specified condition of fixation; each entry of $\eta_{\Delta R}$ is based on the mean value of the corresponding constant errors of the respective groups of 20 settings. The method of computation used is given in the previous experiment, (Lit, 1959).

The effect of conditions of fixation on the magnitude of the stereoscopic threshold angle is shown graphically

in Fig. 1, where log $\eta_{A.D.}$ is plotted as a function of illuminance level, log E. The curves representing the three conditions of fixation are based on the combined data of both *O*s.

For each of the curves shown in Fig. 1, the magnitude of $\eta_{A.D.}$ is largest at the lowest illuminance level and progressively decreases to about the same final low value (9 seconds of arc) as illuminance level is increased. Furthermore, each of the three curves exhibits a marked discontinuity at about the same level of illuminance (− 1.25 log trolands). Also to be noted is the vertical separation of the three curves at all but the highest illuminance levels. The corresponding values of $\eta_{A.D.}$ are characteristically smallest when fixation is maintained on the movable comparison rod (Cond. 1) and characteristically largest when fixation is maintained on the stationary standard rod (Cond. 2). Alternating fixation from comparison to standard rod (Cond. 3) generally yields intermediate values of $\eta_{A.D.}$. Finally, it should be observed that, as illuminance level is decreased, $\eta_{A.D.}$ undergoes about a 19-fold over-all decrease under Cond. 1 and about a 25-fold over-all decrease under Cond. 2 and 3.

The effect of conditions of fixation on the angular magnitude of the corresponding constant errors, $\eta_{\Delta R}$, is more complex and differs markedly for the two *O*s. As in the case of the previous experiment (Lit, 1959), the data of the present study show a characteristic individual difference with respect to the sign of $\eta_{\Delta R}$. Regardless of the condition of fixation used, Observer R. B. consistently adjusts the comparison rod to a position nearer than that of the stationary standard rod, whereas Observer F. C. consistently adjusts the comparison rod to a more distant position than that of the standard rod. When the experimental data for F. C. and R. B. were plotted separately in terms of log $|\eta_{\Delta R}|$ as a function of log E, the obtained curves (not shown here) revealed considerable individual differences. The curves for F. C. showed that for all three conditions of fixation, the magnitude of $\eta_{\Delta R}$ is large at low levels of illuminance and progressively decreases to about the same final value (about 50 seconds of arc) as illuminance level is increased. All three curves for F. C. revealed a discontinuity in the function at about the same illuminance level (− 1.25 log trolands). At all but the highest illuminance levels, the corresponding values of $\eta_{\Delta R}$ for this *O* were smallest for fixation on the comparison rod (Cond. 2) and largest for alternating fixation (Cond. 3). It was also noted for F. C. that $\eta_{\Delta R}$ undergoes about an eight-fold over-all decrease for Cond. 1, and about a ten-fold over-all decrease for Cond. 2 and 3. The corresponding data for R. B. were considerably more irregular, particularly at low illuminance levels (below − 1.25 log troland) for alternating fixation (Cond. 3). At these low illuminance levels, the corresponding values of $|\eta_{\Delta R}|$ were observed to be smallest for Cond. 1; for Cond. 2 and 3 the corresponding values of $|\eta_{\Delta R}|$ were larger and about equal. For illuminance levels above − 1.25 log trolands, it was observed that $|\eta_{\Delta R}|$ progressively decreased in magnitude as illuminance level was increased, but the final limiting values of $|\eta_{\Delta R}|$ varied with condition of fixation. Thus $|\eta_{\Delta R}|$ levelled off to a value of about 20 seconds of arc for fixation on the stationary standard rod (Cond. 2) and leveled off to a value of 30 seconds of arc for Cond. 1 and 3.

Discussion

The results for all three conditions of fixation show that illuminance level has an important influence on the magnitude of the stereoscopic threshold angle, $\eta_{A.D.}$; the threshold values are largest (about 200 seconds of arc on the average) at the lowest illuminance level, and progressively decrease to the same final low value (about 9 seconds of arc on the average) as illuminance level is increased by a factor of 100,000. A discontinuity in the rate of decrease of $\eta_{A.D.}$ occurs for all conditions of fixation at an illuminance level of − 1.25 log trolands where $\eta_{A.D.}$ is about 80 seconds of arc on the average. The discontinuity obtained at this illuminance level is typical for many other classes of visual functions and presumably is caused by the transition from rod to cone vision in accordance with duplicity theory (Hecht, 1937).

The data of Mueller and Lloyd (1948) on equality settings in a stereoscope showed a similar discontinuity in the threshold-illuminance function. The results of the present study and those of Mueller and Lloyd do, however, differ considerably with respect to the over-all range of variation in stereoscopic threshold angle. The present study shows that the average value of $\eta_{A.D.}$, for all three conditions of fixation, undergoes more than a 20-fold over-all decrease in magnitude as illuminance level is increased. In contrast, only a three-fold over-all decrease is reported by Mueller and Lloyd. It does not seem likely that this discrepancy between the results of the two studies can be attributed to differences in conditions of fixation. Other stimulus factors that are different in the two experiments are probably responsible.

The data of the present study show that the condition of fixation is an important factor in determining the shape of the threshold-illuminance function. The photopic sections of the curves in Fig. 1 for Cond. 1 and 3 are essentially similar in shape, but the photopic section of the curve for fixation Cond. 1 (fixation on the comparison rod) shows that $\eta_{A.D.}$ decreases at a relatively faster rate, that is, the decrease from cone threshold value to asymptotic threshold value occurs within a relatively smaller range of illuminance levels. Expressed in terms of Hecht's (1937) photochemical equations of intensity discrimination, the best fitting curves for the photopic data of Cond. 2 and 3 are obtained by the theoretical equation which assumes underlying activity of bimolecular chemical reactions, where $m = n = 2$; the best fitting curve for the photopic data of Cond. 1 is obtained by the equation which assumes underlying activity of monomolecular chemical reactions, where $m = n = 1$. The scotopic sections of the curves in Fig. 1, for all three conditions of fixation, can be fitted equally well by either the monomolecular or bimolecular theoretical equations. The variations in the shape of the threshold-illuminance curves of Fig. 1 are not accounted for convincingly by the assumption that the chemical orders of the underlying reactions vary with condition of fixation. Additional experiments on depth discrimination thresholds at many illuminance levels are required to demonstrate that concomitant variations in other factors (for example, accommodation [Alpern, 1958] or convergence) could produce differential effects on the variability of depth settings, depending on the method of fixation used at each of the given illuminance levels.

The data of the present study are in general agreement with the findings of previous investigators (Wright, 1951; Rady & Ishak, 1955; Ogle, 1956) in demonstrating that variations in conditions of fixation produce systematic differences in the magnitude of the corresponding stereoscopic threshold angles. The present data confirm the findings of these previous investigators: for a given lateral separation of the stimulus targets, alternating fixation (Cond. 3) yields a relatively smaller stereoscopic threshold angle than does constant fixation (Cond. 2). The present data show further that a difference in threshold magnitude exists between these two methods of fixation (Cond. 2 and 3) at all but the highest illuminance levels. An even larger difference in corresponding threshold angles could be expected if the lateral separation of the targets used in the present study had been greater than 0.75°.

The present study also provides data on the effects of an additional condition of fixation (Cond. 1). These new data show that, at all but the highest illuminance levels, fixation maintained only on the moving comparison rod yields stereoscopic threshold angles that are generally smaller than the corresponding threshold angles obtained under the other two fixation conditions. This new finding cannot be accounted for merely in terms of differences in convergence effects. Likely, the lower threshold values obtained for Cond. 1 are a result of the fixated comparison target being continuously imaged in each eye on the retinal region having greatest acuity, the foveal region. In any event, the resolu-

tion of the present controversy concerning the relative contributions of stereoscopic parallax and convergence cues to stereoscopic acuity must await additional experiments. Data are required on equality settings obtained at various illuminance levels under each of the three specified conditions of fixation used in the present study for stimulus targets presented over a wide range of observation distances at various lateral separations.

Summary

In this study the precision of depth discrimination has been measured in a two rod test apparatus involving real-depth cues. Effects due to different methods of fixating the rods have been assessed at each of a number of levels of equal illuminance presented to the two eyes. It was found that for all three methods of fixation, the precision of depth discrimination increased by about 20:1 as the level of illuminance was increased over a range of some five log units.

One method of fixation (Cond. 1) involved steady fixation on the movable comparison rod. A second method of fixation (Cond. 2) involved steady fixation on the immovable standard rod. In the third method of fixation (Cond. 3), the *Os* were allowed to fixate on either of the rods at will.

The precision of depth discrimination was clearly greatest under Cond. 1, at all but the very highest illuminance levels at which all three conditions gave essentially equivalent results. Cond. 2 produced generally the lowest precision, whereas Cond. 3 produced levels of precision generally intermediate between the other two.

These results have a significance for photochemical theories of vision and for the current controversy concerning the role of convergence cues in stereoscopic acuity.

References

Alpern, M. Variability of accommodation during steady fixation at various levels of illuminance. *J. Opt. Soc. Amer.*, 1958, **48**, 193–197.

Hecht, S. Rods, cones, and the chemical basis of vision. *Physiol. Rev.*, 1937, **17**, 239–290.

Lit, A. Depth-discrimination thresholds as a function of binocular differences of retinal illuminance at scotopic and photopic levels. *J. Opt. Soc. Amer.*, 1959, **49**, 746–752.

Lit, A., & Hyman, A. The magnitude of the Pulfrich stereophenomenon as a function of distance of observation. *Amer. J. Optom.*, 1951, **28**, 564–580.

Mueller, C. G., & Lloyd, V. V. Stereoscopic acuity for various levels of illumination. *Proc. Nat. Acad. Sci., Wash.*, 1948, **34**, 223–227.

Ogle, K. N. Stereoscopic acuity and the role of convergence. *J. Opt. Soc. Amer.*, 1956, **46**, 269–273.

Rady, A. A., & Ishak, I. G. H. Relative contributions of disparity and convergence to stereoscopic acuity. *J. Opt. Soc. Amer.*, 1955, **45**, 530–534.

Wright, W. D. The role of convergence in stereoscopic vision. *Proc. phys. Soc. Lond.*, 1951, **B64**, 289–297.

(Received January 27, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

ESCAPE LEARNING AS A FUNCTION OF AMOUNT OF SHOCK REDUCTION[1]

G. H. BOWER, H. FOWLER, AND M. A. TRAPOLD[2]

Yale University

Several studies (Crespi, 1942; Zeaman, 1949) have shown that performance in instrumental appetitive conditioning is directly related to the magnitude of reward contingent upon the response. Moreover, these studies have indicated that shifts in amount of reinforcement (amount of food or water) lead to rapid and appropriate changes in performance level.

The results of a study by Campbell and Kraeling (1953) provide evidence that magnitude of reward is an important variable affecting performance in instrumental escape conditioning. Rats were run in a straight alley in which the startbox and runway provided a continuous 400-v. shock and the goalbox either 300-, 200-, 100-, or 0-v. shocks. Running speed was faster the greater the reduction in shock intensity at the goal.

The present study sought to extend these findings on the role of magnitude of reinforcement in escape learning. Specifically, the purpose was to determine whether shifts in the amount of shock reduction would produce changes in escape performance similar to those found in appetitive conditioning with shifts in the amount of goal substance. This question was assessed by training rats initially at one level of shock reduction and then shifting them to different levels during a subsequent testing phase.

[1] This experiment was performed while G. H. Bower and H. Fowler held predoctoral fellowships from the National Institutes of Mental Health.

[2] Now at the State University of Iowa.

METHOD

Subjects.—The *S*s were 35 male, Sprague-Dawley rats, approximately 120 days old. The *S*s had previous experience (30 days prior to testing) in an unrelated experiment on avoidance learning in a bar-pressing situation. Food and water were continuously available in *S*s' living quarters.

Apparatus.—A straight alley, 68 in. long, 3 in. wide, and 4 in. high, served as the test apparatus. The alley was constructed of two L-shaped strips of galvanized sheet metal and was covered with a transparent, Plexiglas top. One of the sheet metal strips formed one side and half of the floor of the apparatus; the other strip formed the other side and other half of the floor. A 1-in. gap separated the two strips at the floor.

The alley was divided into three sections: a startbox 8 in. long, a middle section 48 in. long, and a goalbox 12 in. long. Raising the startbox door operated a microswitch which closed a shock circuit to the alley and started a Standard Electric timer. The clock stopped when *S* broke a photobeam 2 in. inside the goalbox. Running times were recorded to the nearest .01 sec. and were converted to speed scores (1/time).

Two identical but independent shock systems were employed. One provided shock for the startbox and middle sections of the alley, and the other for the goalbox. In each shock system the galvanized metal plates forming the halves of the alley were connected across the output of an AC transformer through a voltage divider which provided variation in the applied voltage. In addition, a .25 meg-ohm resistor was placed in series with *S*. The *S* received shock as long as its feet touched the two strips of sheet metal forming the alley. The narrowness of the alley prevented *S* from avoiding shock by running along only one side of the alley.

Procedure.—The *S* was placed in the startbox with the startbox door closed and the goal box door open. When *S* oriented toward the goal, the startbox door was raised, turning on the shock, and *S* could run to the goal end of the alley. When *S* entered the goalbox and broke the photobeam, which stopped the

clock, the goalbox door was lowered. The *S* was left in the goalbox with the appropriate goal shock for a period of 20 sec. Each *S* received a total of 30 trials. The trials were massed; immediately following removal from the goalbox, *S* was reinserted into the startbox for the next trial.

Experimental design.—The 35 *S*s were randomly assigned to seven groups of five *S*s each. For all groups, and over the entire 30 trials, the shock in the startbox and middle section of the alley was 250 v. The shock in the goalbox was varied across groups and was either 50, 150, or 200 v. After Trial 15, the goal shock for four of the seven groups was shifted from one to another of these voltages. The goal shock voltage for each group on Trials 1–15 and 16–30 is indicated in Table 1.

Results

Speed performance curves are presented in Fig. 1. The group median speed over blocks of three trials was used to yield smooth curves. All statistical analyses were performed on group mean speeds over Trials 10–15 or 25–30.

Group mean speeds for these trials are presented in Table 1. The score to the left in each cell is for Trials 10–15; the score to the right is for Trials 25–30. As may be seen in Table 1, the three subgroups initially on the 50-v. condition did not differ on Trials 10–15 ($F < 1$); similarly, the subgroups initially on the 200-v. condition during Trials 1–15 did not differ on Trials 10–15 ($F < 1$). These

TABLE 1

Group Mean Speeds (1/Time) on Trials 10–15 and 25–30

Goal Shock on Trials 1–15	Goal-shock on Trials 16–30					
	50		150		200	
	Tr. 10–15	Tr. 25–30	Tr. 10–15	Tr. 25–30	Tr. 10–15	Tr. 25–30
50	.65	.69	.67	.43	.65	.25
150			.43	.43		
200	.20	.67	.17	.41	.11	.22

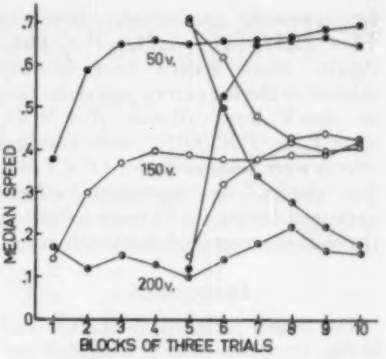

Fig. 1. Group median speed as a function of goal-shock voltage. See text for further explanation.

results indicate that the randomization was effective and that the subgroups within a given condition were equated prior to the shift in goal shock after Trial 15. Thus, these subgroups are pooled in all relevant comparisons over Trials 1–15.

The acquisition curves for groups receiving 50, 150, or 200 v. in the goalbox during Trials 1–15 are shown in the left half of Fig. 1. The curves diverged over training, and reached different asymptotes. The differences in asymptotic performance (Trials 10–15) are statistically significant ($F = 29.02$, $df = 2,12$, $P < .001$). The asymptotic mean speeds were linearly related to the log percentage reduction in shock (linear $F = 57.34$, $df = 1,12$, $P < .001$); residual component effects were not significant ($F < 1.00$).

As may be seen in the right half of Fig. 1, the groups shifted in goal-shock voltage showed rapid changes in speed, approaching the appropriate asymptotic levels of performance of the nonshifted groups. By Trials 25–30, the effect of initial training condition had disappeared ($F < 1.00$) and performance was determined solely

by present goal-shock intensity ($F = 25.52$, $df = 2,24$, $P < .001$). Again, mean speeds were linearly related to the log percentage reduction in shock level (linear $F = 50.47$, $df = 1,24$, $P < .001$); and residual effects were nonsignificant ($F < 1.00$). No "elation" or "depression" effects appeared during the 15 trials following the shift in amount of shock reduction.

Discussion

The results presented in the left half of Fig. 1 confirm those of Campbell and Kraeling (1953) in showing faster escape performance for larger amounts of shock reduction. The magnitude of reward appears to be linearly related to the log percentage reduction in shock level over the range of values studied. This conclusion is essentially that reported by Campbell and Kraeling.

The results presented in the right half of Fig. 1 show that appropriate and rapid changes in performance occur following shifts in the amount of shock reduction. These results parallel those of Crespi (1942) and Zeaman (1949) obtained with appetitive rewards and suggest that similar molar laws operate for magnitude of reward in both instrumental appetitive and escape conditioning.

In the context of the behavior theory developed by Hull (1952) and Spence (1956), the results of this study suggest that magnitude of shock reduction can be handled by an intervening variable which bears the same relation to habit strength and drive as does the incentive motivation variable (K) in instrumental appetitive conditioning. Thus, the diverging curves in Fig. 1 indicate that the magnitude of shock reduction combines multiplicatively with trials (N) to determine speed. The performance shifts following changes in the magnitude of shock reduction are also consistent with this formulation. In instrumental appetitive conditioning the incentive motivation variable is closely related to the $r_g - s_g$ mechanism (Spence, 1956). However, no analogous mechanism is obvious in the case of escape conditioning.

Summary

Rats ran from an alley where they received a continuous shock of 250 v. to a goalbox where they received a continuous shock of either 50, 150, or 200 v. for 20 sec. The amount of shock reduction on Trials 1–15 and Trials 16–30 was varied according to an incomplete factorial design. The results showed that escape performance was an increasing function of amount of shock reduction, and that shifts in amount of shock reduction led to rapid and appropriate changes in level of performance. The results are analogous to those obtained by Crespi (1942) and Zeaman (1949) with appetitive rewards and suggest that similar behavioral laws underlie both instrumental appetitive and escape conditioning.

References

Campbell, B. A., & Kraeling, D. Response strength as a function of drive level and amount of drive reduction. *J. exp. Psychol.*, 1953, **45**, 97–101.

Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.

Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

Zeaman, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

(Received January 28, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

SPATIAL CONTIGUITY OF CUE, REWARD, AND RESPONSE IN DISCRIMINATION LEARNING BY CHILDREN[1]

J. V. MURPHY[2] AND R. E. MILLER

Department of Clinical Science, School of Medicine, University of Pittsburgh

Discrimination learning has been historically one of the most important methods, along with maze learning and conditioning, for the investigation of learning phenomena. There have been numerous experiments dealing with the nature of the discriminative stimulus, and with characteristics of the reward and its presentation. Many significant contributions concerning the learning process and its parameters have resulted from the discrimination paradigm.

Recently a new facet of the learning process has been studied by the discrimination method. Several investigations of the spatial relationship between the cue, response, and reward in learning have been reported (Gardner & Nissen, 1948; McClearn & Harlow, 1954; Murphy & Miller, 1955; Murphy & Miller, 1958). A relatively small spatial disparity between the cue and the site of the reward and response virtually obviates learning in infrahuman primates (Murphy & Miller, 1955; Murphy & Miller, 1958). These same species, however, can learn conventional discriminations readily despite rather extensive cortical lesions (Moss & Harlow, 1948) or gross differences in motivational levels (Meyer, 1951). It is, therefore, rather surprising that the lack of spatial contiguity has proven to be such a sensitive variable in discrimination learning.

The "spatial phenomenon" was examined in more detail in a recent investigation in which a standard spatial distance (7 in.) was introduced between cue, response, and reward in each of the five possible combinations, i.e., cue and response separated from reward, cue and reward separated from response, etc. (Murphy & Miller, 1958). It was found that cue disparity was the critical factor in the "spatial phenomenon." So long as the cue was spatially contiguous with either the response or the reward, learning proceeded normally; however, learning was significantly retarded when the cue was separated from both response and reward.

The present study was designed to investigate the importance of this spatial factor in man. Specifically, children at several grade/age levels were presented with a two-choice discrimination problem. The experiment was designed to determine the effect of spatial disparities between the cue and the rewarded response upon both learning and transfer.

METHOD

Subjects.—Forty-one children from a suburban community were randomly selected from school Grades 1 through 4. The only restrictions were that they had not been held back in grade advancement and that there were relatively equal numbers of boys and girls.

Apparatus.—The apparatus, as shown in Fig. 1, was a sliding formboard and a vertical

[1] This research was supported by a research grant (M-487 C5) from the National Institute of Mental Health of the National Institutes of Health, United States Public Health Service. We are indebted to the teachers and administrators of the Penn Hills School System, Penn Hills, Pa., for the excellent cooperation extended throughout the course of the study.

[2] Present address: Systems Development Corporation, Santa Monica, California.

(A) (B)

FIG. 1. The discrimination apparatus and formboard showing: (A) the nonseparated condition; and (B) the cue-separated condition.

partition containing a door which made it possible to set up the problem out of *S*'s view. The formboard was designed to pass under the open screen with only ¼-in. clearance in order to preclude the observation of *E* by *S* and vice versa. The operative movements of this apparatus were identical to a modified form of the Wisconsin General Test Apparatus (WGTA) employed in similar work with monkeys (Murphy & Miller, 1958). The formboard itself resembled those used in previous studies in the spatial area (McClearn & Harlow, 1954; Murphy & Miller, 1958). A unique feature was an attempt to obscure the importance of lines running between the cue and the site of the reward and response. To accomplish this purpose, the surface was cut with slots some ¼-in. deep to simulate those which occurred around the stimulus blocks. The 2-in. square stimuli were painted on 2-in. × 8-in. blocks which were removable. The stimulus blocks had a plywood backing which extended ¼ in. on all sides. This backing eliminated all light cues from around the stimulus blocks. A 2-in. wide board affixed to the back of the formboard upright just above the response blocks assured that the stimulus blocks did not move if bumped when the response blocks were moved. A magnetic cabinet latch was recessed into the top above each stimulus block. A metal clip on the plywood backing of each stimulus block served to hold the top of the stimulus blocks securely in place. The final result of these features was two stimulus blocks which could be readily removed and interchanged in position with a minimum of effort and artifactual cues, and with a maximum degree of rigidity when in place. The stimulus blocks were attached to the plywood backing by a centrally located screw. Thus, to introduce the spatial factor it was necessary only to rotate the stimulus blocks and reinsert a locating pin. The stimuli were located 12 in. apart. The response blocks were made of ¾-in. hardwood with a clear varnish finish. They were 2-in. wide and extended 1½-in. in front of the stimulus blocks. Stops attached to the sides of the response blocks in back of the upright portion ensured that movement was possible in one direction only, i.e., away from *S*. A ¼-in. × 1-in. channel was cut into the base of the formboard extending from behind the upright portion to the foodwell. The channels and the foodwells both contained a cushioning layer of masking tape. These features made it possible to place the reward conveniently and noiselessly in the foodwell by placing the reward in the channel on *E*'s side of the formboard. The reward was then pushed into the foodwell as the response blocks were inserted for the next trial.

Procedure.—The two conditions of the discrimination problem are shown in Fig. 1. The nonseparated condition (Fig. 1A) is the conventional discrimination situation in which cue, response, and reward are all spatially contiguous. The *S* must learn to respond by pushing a response block located beneath the positive stimulus (in this case the stimulus on which the base of the white triangle is down) in order to secure a reward concealed beneath the response block to solve the problem. The same solution is required in the cue-separated condition (Fig. 1B). In this case, however, the cue is located 6 in. above the response block. The response and

reward are contiguous in the cue-separated condition.

The tests were conducted during regularly scheduled school hours in an unused basement area where distractions were minimal. Each *S* was taken individually to the testing room and seated before the apparatus. The following instructions were given: "I want you to try to find the candy. It is always under this block (indicated) or under this one." (The block under the positive stimulus was displaced and *S* was allowed to take the candy reward.) It was then suggested that some candy might be saved for later, and a closed container was provided as a receptacle. The reward consisted of one piece of sugar-coated chocolate candy. In the test situation the opaque screen was lowered while the reward was appropriately positioned. Both stimulus blocks were replaced in the apparatus at the same time to eliminate the possibility of providing extraneous cues. The same procedure was followed with the response blocks. After the problem had been prepared, the opaque screen was raised and the stimulus tray pushed within reach of *S*. If *S* made a correct choice the reward was immediately apparent in the foodwell located beneath the response block. A noncorrection procedure was used throughout. The position of the positive stimulus was varied from side to side in accordance with the Gellerman series (Gellerman, 1933). Each *S* was given 50 trials followed by a 3-min. rest period. The rest period was occupied by going upstairs for a drink of water. An additional 50 trials were then administered for a total of 100 trials a day.

The experimental design was to test one group under the nonseparated condition and to test a comparable group under the cue-separated condition. The spatial factor was also introduced for *S*s who had attained an acquisition criterion under the nonseparated condition, in order to obtain information regarding transfer of the learned discrimination to the spatially separated condition.

Since the ability level of humans on this particular learning problem was unknown, it was decided to initiate testing under the nonseparated condition in Grade 1. The *S*s were continued for a maximum of 400 trials under the nonseparated condition or to a criterion of 10 successive errorless trials. After attaining the criterion, the testing was terminated for that day; on the following day *S* was tested on the cue-separated condition until 10 successive errorless trials had been attained, but with a maximum of 100 trials. Three *S*s were tested each week. Successive grades were tested until the level was determined at which more than 50% of *S*s were able to solve the nonseparated arrangement of the problem within the maximum of 400

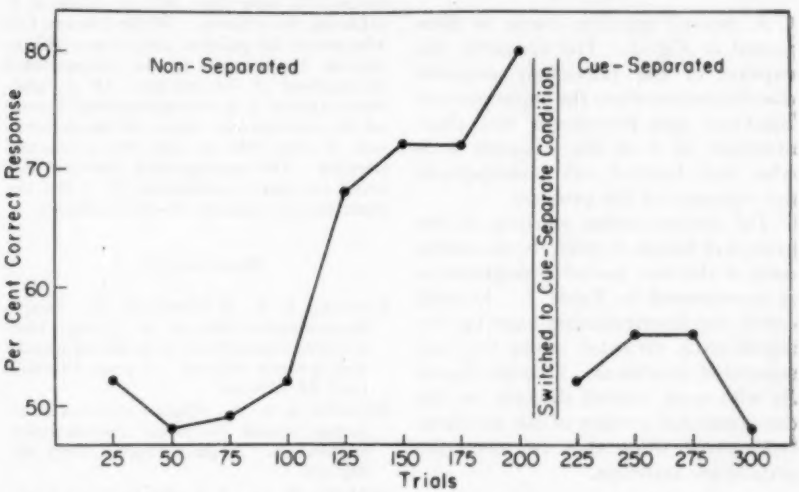

FIG. 2. The learning of the discrimination under the nonseparated condition by a typical *S* and subsequent transfer to the cue-separated condition.

TABLE 1

THE PERFORMANCE OF GRADE 4 CHILDREN IN DISCRIMINATION LEARNING

Condition	*N*	Mean Errors	*SD*	*t*
Cue-separated	9	167	31	3.5*
Nonseparated	9	92	52	

* $P < .01$.

trials. The data collected from *Ss* from Grade 3 suggested that *Ss* from Grade 4 would be able to solve the nonseparated version of the problem within 400 trials. Consequently, the first nine *Ss* from Grade 4 were tested on the cue-separated arrangement of the problem. Nine additional *Ss* were then tested on the nonseparated arrangement, and were switched to the cue-separated arrangement on the day after they attained the criterion in the nonseparated condition. When *S* had completed the experimental regimen an additional 30 trials were administered in which the problem was scrambled so that the former solution was incorrect and the problem unsolvable. At this point the somewhat confused *S* was asked not to tell his friends how to solve the problem.

RESULTS AND DISCUSSION

A typical learning curve is illustrated in Fig. 2. The dramatic disruption of the previously acquired discrimination when the cue-separated condition was introduced was characteristic of 5 of the 9 Grade 4 *Ss* who first learned the nonseparated arrangement of the problem.

The discrimination learning of two groups of Grade 4 children, one under each of the two spatial arrangements, is summarized in Table 1. As indicated, the discrimination learning was significantly retarded under the cue-separated condition. Of nine Grade *Ss* who were started directly on the cue-separated version of the problem, only three were able to meet the acquisition criterion.

These results are very much in accord with previous data from monkeys (Murphy & Miller: 1955, 1958) with regard to both learning and transfer.

Since the results show that the spatial factor is also an important variable in discrimination learning in man, it is suggested that some attention to the spatial relations in learning situations would faciliate learning or perhaps permit learning in otherwise unlearnable situations. The extent to which man can learn to overcome spatial disparities and the developmental trends of such spatial facility are unknown. It can be asserted, however, that displacement of stimuli from the site of the reward and response is a very potent variable in the discrimination learning of both man and the infrahuman primates.

SUMMARY

An experiment was designed to investigate the effect of separating the cue from the reward and response on the discrimination learning of elementary school children. Thirty-two *Ss* in Grades 1 through 4 were tested on the nonseparated arrangement of the problem. Performance on the discrimination problem with the cue, reward, and response contiguous improved at successive grade levels with more than 50% of Grade 4 *Ss* achieving the criterion. Of the 8 Grade 4 *Ss* who solved the problem, only 3 were able to transfer the solution to the cue-separated arrangement of the problem. Of an additional 9 Grade 4 *Ss* who were started directly on the cue-separated version of the problem, only 3 were able to meet the acquisition criterion. The cue-separated condition was found to retard significantly ($P < .01$) the discrimination learning of school children.

REFERENCES

GARDNER, L. P., & NISSEN, H. W. Simple discriminative behavior of young chimpanzees: Comparisons with human aments and domestic animals. *J. genet. Psychol.*, 1948, **72**, 145–164.

GELLERMAN, L. W. Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, 1933, **42**, 206–208.

JENKINS, W. O. A spatial factor in chimpanzee learning. *J. comp. physiol. Psychol.*, 1943, **35**, 81–84.

McCLEARN, G. E., & HARLOW, H. F. The effect of spatial contiguity on discrimination learning by rhesus monkeys. *J. comp. physiol. Psychol.*, 1954, **47**, 391–394.

MEYER, D. R. Food deprivation and discrimination reversal learning by monkeys. *J. exp. Psychol.*, 1951, **41**, 10–16.

Moss, E., & HARLOW, H. F. Problem solution by monkeys following extensive unilateral decortication and prefrontal lobotomy of the contralateral side. *J. Psychol.*, 1948, **25**, 223–226.

MURPHY, J. V., & MILLER, R. E. The effect of spatial contiguity of cue and reward in the object-quality learning of rhesus monkeys. *J. comp. physiol. Psychol.*, 1955, **48**, 221–229.

MURPHY, J. V., & MILLER, R. E. The effect of the spatial relationship between the cue, reward, and response in simple discrimination learning. *J. exp. Psychol.*, 1958, **56**, 26–31.

(Received February 2, 1959)

Journal of Experimental Psychology
Vol. 58, No. 6, 1959

SUPPLEMENTARY REPORT: STIMULUS AND RESPONSE MEANINGFULNESS (m') IN PAIRED-ASSOCIATE LEARNING BY HOSPITALIZED[1] MENTAL PATIENTS

VICTOR J. CIEUTAT
Louisiana State University

This research was undertaken as an extension of two experiments by Cieutat, Stockwell, and Noble (1958). The *Ss* were 24 male patients, whose hospital diagnoses were predominantly schizophrenic reactions, randomly selected from the VA Continued Treatment Service. Ages ranged from 24 to 69 yr. (mean = 42.5). Six were rejected because of inability to cooperate. Each *S* received 10 trials on a practice list of 4 pairs of CVC syllables, a 2-min. rest, and 30 trials on one of the four experimental lists. The mean m' values of the S and R terms in the latter lists varied from low (L = mean m' of 0.68) to high (H = mean m' of 3.02) in all possible combinations. Three random trial sequences were prepared for pairs in the experimental lists. Two *Ss* began their learning on each of the trial sequences within each list, permitting evaluation of the order in which the sequences were learned. Standard instructions were given. Terms were spelled and the correction method was used.

Results.—The median test indicated *Ss* in the four groups did not significantly differ in initial learning ability as measured by number of correct anticipations ($\Sigma R+$) on the practice list ($\chi^2 = 0.53$; $df = 3$; $P > .05$).

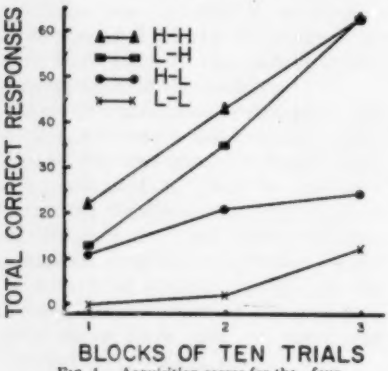

FIG. 1. Acquisition scores for the four experimental lists.

TABLE 1

GROUP MEANS ACROSS 30 TRIALS

m' for S Term	m' for R Term		Row Means
	L (0.68)	H (3.02)	
L (0.68)	2.33	18.50	10.42
H (3.02)	9.33	21.33	15.33
Means	5.83	19.92	

Two levels of both S and R meaningfulness, three orders of trial sequences, and three 10-trial blocks of trials were evaluated in a $2 \times 2 \times 3 \times 3$ mixed analysis of variance. Only the effect of trials was significantly related to variations in $\Sigma R+$ on the experimental lists ($F = 10.42$; $df = 2/24$; $P < .01$). Acquisition scores of the four groups, disregarding order, are given in Fig. 1, and group means across 30 trials are in Table 1.

Discussion.—These results are interpreted as an extension of previous findings because: (*a*) the largest simple effect is that between the H-H and L-L groups, as found with analogous groups in prior results; (*b*) the mean difference on the R side is greater than that on the S side (14.09 vs. 4.91), also identical with the prior findings; and (*c*) the rank order of the group means is exactly that found for similar groups in both earlier experiments. It is thus considered more reasonable to attribute the failure of these effects to attain statistical significance to the small *n*, and the consequent lack of precision, than to the absence of these functions in the present population. It is therefore concluded that these results constitute an extension of earlier findings to a population of hospitalized mental patients, illustrating the within-species generality of certain functions in the area of human learning.

REFERENCE

CIEUTAT, V. J., STOCKWELL, F. E., & NOBLE, C. E. The interaction of ability and amount of practice with stimulus and response meaningfulness (m, m') in paired-associate learning. *J. exp. Psychol.*, 1958, 56, 193–202.

(Received March 28, 1959)

[1] This research was done while the writer was a Clinical Trainee at the VA Hospital, Gulfport, Mississippi, 1958.

ZEITSCHRIFT FÜR
EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Journal of the German Psychological Association

Editors for Germany

PROF. DR. PH. LERSCH, München; PROF. DR. F. SANDER, Bonn

Editor for Austria
PROF. DR. H. ROHRACHER, Vienna

Editor for Switzerland
PROF. DR. R. MEILI, Bern

Redaction
DR. C. J. HOGREFE, Göttingen

Volume 6, No. 1

CHARLOTTE BÜHLER. **General Fundamentals and Problems of Modern Psychotherapy**

HILDEGARD HETZER. **The Body in Self-Portrayals by Children in the Year 1926 and in the Year 1957**

PETER R. HOFSTÄTTER. **Self-Knowledge**

FRIEDRICH KAINS. **"Alalia ex Separatione"**

PAUL F. LAZARSFELD. **American Observations on a Student of Bühler**

JEAN PIAGET. **The Relational Method in the Psychology of Perception**

HUBERT ROHRACHER. **Psychological Problems of Control**

ALBERT WELLEK. **A Third of a Century after Bühler's "Crisis in Psychology"**

KONRAD LORENZ. **Form Perception as a Source of Scientific Knowledge**

All articles are provided with summaries in English and French.

The journal appears quarterly and contains 160 pages per issue.
Subscription: DM 64.—per volume.

VERLAG FÜR PSYCHOLOGIE—DR. C. J. HOGREFE—GÖTTINGEN
Brentanoweg 10, Germany

To Be Published February 1960

PLANS AND THE STRUCTURE OF BEHAVIOR

George A. Miller, Harvard Univ.
Eugene Galanter, Univ. of Pennsylvania
Karl H. Pribram, Stanford Univ.

In the past quarter century ideas developed in communication and computer theory have been slowly modifying the way behavioral scientists view living systems and thus enriching our conception of what a human mechanism can be and do. In *Plans and the Structure of Behavior* these changes are made explicit in psychological terms. The basic "unit" affecting human behavior considered in this book is not the S-R connection, but the feed-back loop. New dimensions are added to traditional psychological topics when the organism is viewed as an information processing system. The neurological basis of the processes is discussed throughout.

$4.75 (probable price)

PSYCHOLOGY OF PERCEPTION

William N. Dember
University of Cincinnati

A psychological approach to data, method, and interpretation in the study of perception. *Contents:* Introduction * Threshold Measurement Techniques * Attributes, Similarity, and Stimulus Location * Visual Psychophysics * The Organization of Visual Perception * The Influence of Context * Effects of Learning on Perception * Set and Perception * Motivational Effects on Perception * Attention, Motivation, and Emotion.

$6.00 (probable price)

Outstanding 1959 Publications

PSYCHOLOGY AND HUMAN PERFORMANCE: AN INTRODUCTION TO PSYCHOLOGY

Robert M. Gagne, Princeton Univ.
Edwin A. Fleishman, Yale Univ.

An experimentally oriented course in the principles of psychology. Stress on their applications leads to an understanding of human behavior. *$7.25*

WORKBOOK FOR PSYCHOLOGY AND HUMAN PERFORMANCE

W. F. Strother, Princeton Univ.
R. L. Decker, West Virginia Univ.

Contains self-test exercises, quizzes, and problems coordinated with the text. Includes "How to Study" chapter.

$1.50 paper

APPLICATIONS OF INFORMATION THEORY TO PSYCHOLOGY

Fred Attneave
University of Oregon

Gives the reader with no knowledge of information theory the understanding he needs in order to apply the techniques of informational statistics to psychological research. The author provides an excellent summary of the uses of information theory in psychology and a discussion of the results of such applications. *$3.75*

HENRY HOLT AND COMPANY, INC., 383 Madison Ave., N.Y.